

Uniform Diameter Bounds in Branch Groups

Henry Bradford

March 20, 2017

Abstract

Let G be either the Grigorchuk 2-group or one of the Gupta-Sidki p -groups. We give new upper bounds for the diameters of the quotients of G by its level stabilisers, as well as other natural sequences of finite-index normal subgroups. Our bounds are independent of the generating set, and are polylogarithmic functions of the group order, with explicit degree. Our proofs utilize a version of the profinite Solovay-Kitaev procedure, the branch structure of G , and in certain cases, existing computations of the lower central series of G .

1 Introduction

Let G be a finite group, and $S \subseteq G$ be a generating set. The *diameter of G with respect to S* is defined to be:

$$\text{diam}(G, S) = \min\{n \in \mathbb{N} : B_S(n) = G\},$$

where $B_S(n)$ is the (closed) ball of radius n about the identity in the word metric defined by S on G . The *diameter of G* , denoted $\text{diam}(G)$, is then the maximal value of $\text{diam}(G, S)$ as S ranges over all generating subsets of G . In this paper we give upper bounds for the diameters of natural families of finite quotients of certain *branch groups*.

1.1 Statement of Results

Theorem 1.1. *Let \mathfrak{G} be the Grigorchuk 2-group. Then:*

$$\begin{aligned} \text{diam}(\mathfrak{G}/\text{Stab}_{\mathfrak{G}}(n)) &= O(\exp(\log(35)n)) \\ &= O(\log|\mathfrak{G} : \text{Stab}_{\mathfrak{G}}(n)|^{\log(35)/\log(2)}). \end{aligned}$$

We shall define the sequence of *level stabilisers* $\text{Stab}_G(n)$ for a group G acting on a rooted tree in subsection 2.2. Our proof makes extensive use of the description of the lower central series $(\gamma_n(\mathfrak{G}))_n$ of \mathfrak{G} given in [5], building on [26] (see also [3]). The results of these papers facilitate an explicit description of the restriction to $\gamma_n(\mathfrak{G})$ of the action of \mathfrak{G} on the binary rooted tree. Indeed, Theorem 1.1 is proved as a consequence of the following.

Theorem 1.2.

$$\begin{aligned} \text{diam}(\mathfrak{G}/\gamma_n(\mathfrak{G})) &= O(n^{\log(35)/\log(2)}) \\ &= O(\log|\mathfrak{G} : \gamma_n(\mathfrak{G})|^{\log(35)/\log(2)}). \end{aligned}$$

Recall that $\log(35)/\log(2) \approx 5.129$. We now turn to our results on the Gupta-Sidki p -groups.

Theorem 1.3. *Let p be an odd prime. Let $\Gamma_{(p)}$ be the Gupta-Sidki p -group. Then:*

$$\begin{aligned} \text{diam}(\Gamma_{(p)}/\text{Stab}_{\Gamma_{(p)}}(n)) &= O_p(\exp(\log(C_p)n)) \\ &= O_p(\log|\Gamma_{(p)} : \text{Stab}_{\Gamma_{(p)}}(n)|^{\log(C_p)/\log(p)}) \end{aligned}$$

where $C_p = 3 \cdot 4^p - 2^p(p+8) + 7$.

Theorem 1.3 is a consequence of our next result. Let K be the derived subgroup of $\Gamma_{(p)}$.

Theorem 1.4. *Let C_p is as in Theorem 1.3.*

$$\begin{aligned} \text{diam}(\Gamma_{(p)}/K^{(\times p^n)}) &= O_p(\exp(\log(C_p)n)) \\ &= O_p(\log|\Gamma_{(p)} : K^{(\times p^n)}|^{\log(C_p)/\log(p)}). \end{aligned}$$

Here $K^{(\times p^n)}$ denotes the Cartesian product of p^n copies of K . We define the natural embeddings of the $K^{(\times p^n)}$ as finite-index normal subgroups of $\Gamma_{(p)}$ in subsection 2.2. For now suffice to say that there are inclusions $K^{(\times p^n)} \leq \text{Stab}_{\Gamma_{(p)}}(n+1)$ and that given certain well-known bounds on the orders of the relevant groups, Theorem 1.3 quickly follows from these inclusions and Theorem 1.4. In the case $p = 3$, we can exploit the description of the lower central series of $\Gamma_{(3)}$ given by Bartholdi [3] to also deduce the following.

Theorem 1.5.

$$\begin{aligned} \text{diam}(\Gamma_{(3)}/\gamma_n(\Gamma_{(3)})) &= O(n^{\log(111)/\log(1+\sqrt{2})}) \\ &= O(\log|\Gamma_{(3)} : \gamma_n(\Gamma_{(3)})|^{\log(111)/\log(3)}). \end{aligned}$$

For the Gupta-Sidki 3-group therefore, our bounds for the diameter are polylogarithmic in the order of the group, with degree $\log(C_3)/\log(3) = \log(111)/\log(3) \approx 4.287$. As p grows, the degree of the polylogarithm grows, proportional to $p/\log(p)$. The implied constants in Theorems 1.1-1.5 may all be explicitly computed from our proofs.

$\mathfrak{G}/\text{Stab}_{\mathfrak{G}}(n)$ and $\Gamma_{(p)}/\text{Stab}_{\Gamma_{(p)}}(n)$ are transitive imprimitive permutation groups on, respectively, 2^n and p^n points. As such, Theorems 1.1 and 1.3 provide new examples of transitive subgroups of $\text{Sym}(N)$ (N a power of a fixed prime) whose diameters are polynomially bounded in N . In the next subsection we will further contextualise Theorems 1.1 and 1.3 within the existing literature on diameters of permutation groups. \mathfrak{G} and $\Gamma_{(p)}$ are particularly famous members of the class of branch groups, and have been extensively studied since their introduction, respectively in [16] and [19]. They will be defined precisely in subsection 2.2.

1.2 Background and Structure of the Paper

\mathfrak{G} is one of the most exotic objects in geometric group theory: it is a finitely generated infinite 2-group, so provides a counterexample to the General Burnside Problem; it is a group of *intermediate growth*, and indeed was the first example of such a group to be constructed; it is amenable but not elementary amenable; it is a residually finite just-infinite group; it admits no faithful representation over any field, and every finite 2-group embeds into it as a subgroup [10]. Similarly, $\Gamma_{(p)}$ is a finitely generated infinite p -group; contains a copy of every finite p -group, and shares many of the other aforementioned properties with \mathfrak{G} , though its growth is the subject of ongoing discussion.

An understanding of the broader class of *branch groups*, to which \mathfrak{G} and $\Gamma_{(p)}$ belong, has now become a crucial part of the toolkit of the modern geometric or profinite group theorist. In one sense this is no surprise, bearing in mind Wilson’s classification of just-infinite groups [29], in which branch groups comprise an important case (although historically, this may only be noted in hindsight, since the class of branch groups was not formally defined until some time after Wilson’s theorem [18]). What was perhaps less expected was the extraordinarily relevance branch groups would prove to have, to subjects as diverse as (but by no means limited to) decision problems, finite automata, spectral graph theory, fractal spaces, and exotic phenomena in the domains of word growth, subgroup growth and other asymptotic invariants of infinite groups [6].

Separate from, but roughly concurrent with these developments, interest was growing in generation problems in various families of groups, including the search for good diameter bounds. From the start, particular attention was paid to upper bounds for permutation groups, owing to connections with problems in theoretical computer science. These included membership testing protocols in computational group theory and complexity analysis of deciding solvability of various combinatorial puzzles [13, 23], the most famous of which is of course the Rubick’s cube. More recently intense interest in diameters of finite groups has been renewed, motivated by connections with such diverse topics as expander graphs, approximate groups, the Banach-Ruziewicz problem, Apollonian circle packings, sum-product phenomena in fields and affine sieves. The modern study of diameters of groups is therefore an extremely rich and diverse subject, and one which we cannot hope to fully capture here; we instead refer the interested reader to [20, 28] and references therein for an overview of the recent developments.

Among the key tools to have featured in proofs of good upper bounds for the diameters of finite groups is the *Solovay-Kitaev procedure*. This is a method which was originally used in the context of compact complex Lie groups (where it had applications to problems in quantum computer science) but which translated readily to the setting of abstract or profinite groups. Given a group Γ and a descending sequence $(\Gamma_i)_i$ of finite-index normal subgroups, the Solovay-Kitaev procedure proves, under suitable additional assumptions, an upper bound on the diameters of the finite quotient groups Γ/Γ_i by induction on the sequence $(\Gamma_i)_i$. A crucial ingredient facilitating the induction step is that any element of a later term in the sequence should be expressible as (or at least sufficiently

closely approximable by) the product of a small number of commutators of elements lying in earlier terms. Exactly what this means in practice will depend on the specific group Γ and sequence of subgroups with which we are working.

In the above setting the Solovay-Kitaev procedure was first used by Gamburd and Shahshahani [14] to give a polylogarithmic upper bound on the diameter of $\mathrm{SL}_2(\mathbb{Z}/p^n\mathbb{Z})$. Their work was subsequently extended, first by Dinai [12] to groups of $(\mathbb{Z}/p^n\mathbb{Z})$ -points of other Chevalley groups, then by the author [8] to congruence quotients of other p -adic analytic groups, $\mathbb{F}_q[[t]]$ -analytic groups and the Nottingham groups of finite fields.

Producing the commutator expressions required to facilitate the induction step in the Solovay-Kitaev procedure requires fairly explicit computations of commutators in the subgroups Γ_i . As such, it is very useful in practice for the sequence $(\Gamma_i)_i$ to be highly “recurrent” in some sense, so that calculations carried out at one level may be translated to others. In many of the examples considered in [8], for instance, the group Γ has a naturally associated Lie algebra over a non-archimedean field, such that the Γ_i may be identified with a descending sequence of balls in the Lie algebra. Under this identification, computations of commutators at different levels are simply “rescalings” of each other.

The regular branch structure of \mathfrak{G} and $\Gamma_{(p)}$ provides a notion of “recurrence” of a different sort. These groups each have a finite-index normal subgroup K which naturally contains a direct product $K^{(\times p)}$ of copies of itself, again as a finite-index normal subgroup (for \mathfrak{G} we take $p = 2$). We may therefore consider the descending sequence $\Gamma_i = K^{(\times p^i)}$, and note that a commutator in Γ_{i+1} is a p -tuple of commutators in Γ_i , thereby translating commutator calculations at the top few levels down to all other levels.

The best previous diameter bounds for quotients of branch groups make no use of their branch structure, or indeed anything other than the fact that they are transitive permutation groups. Babai and Seress obtained the following very general result.

Theorem 1.6 ([1] Theorem 1.4). *Let G be a transitive permutation group of degree N . Then:*

$$\mathrm{diam}(G) \leq \exp(C(\log(N))^3) \mathrm{diam}(\mathrm{Alt}(m(G)))$$

where C is an absolute constant, and $m(G)$ is the maximal degree of an alternating composition factor of G .

If Γ is a regular branch group acting on the k -ary rooted tree, then $G = \Gamma / \mathrm{Stab}_\Gamma(n)$ has a natural transitive action on the n th level of the tree, so Theorem 1.6 is applicable, with $N = k^n$ and $m(G) \leq k$. For $\Gamma = \mathfrak{G}$ or $\Gamma_{(p)}$, $\Gamma / \mathrm{Stab}_\Gamma(n)$ is a p -group of order $\Omega(p^{\Omega(p^n)})$ (see Lemma 2.9 and Corollary 4.1 below), Theorem 1.6 yields:

$$\begin{aligned} \mathrm{diam}(\Gamma / \mathrm{Stab}_\Gamma(n)) &= O\left(\exp\left(O(n^3 \log(p)^3)\right)\right) \\ &= O\left(\exp\left(O_p(\log \log |\Gamma : \mathrm{Stab}_\Gamma(n)|^3)\right)\right) \end{aligned}$$

(here we take $p = 2$ for $\Gamma = \mathfrak{G}$). Theorem 1.6 makes use of some deep machinery, including the Classification of Finite Simple Groups. Therefore our Theorems 1.1-1.5 improve upon prior results in at least three ways:

- (i) They improve the diameter bound qualitatively, from a quasipolynomial function of $\log|G|$ to a polynomial one.
- (ii) They give explicit (and small) estimates for implied constants.
- (iii) They have self-contained, elementary and constructive proofs, which could in principle be implemented in reasonable time on a computer.

It is also noteworthy that Theorem 1.6 remains a key tool in the study of other transitive permutation groups, which are consequently not known to have diameter less than $\exp(C(\log(N))^3)$. The best known bound in this direction is the following result of Helfgott and Seress.

Theorem 1.7 ([21]). *Let G be as in Theorem 1.6. Then:*

$$\text{diam}(G) \leq \exp(C(\log(N))^4 \log \log(N)).$$

By Theorem 1.6, Theorem 1.7 can be immediately reduced to the case $G = \text{Alt}(N)$. The proof in this special case also uses Theorem 1.6 to facilitate an important induction. It is a longstanding conjecture that $\text{Sym}(N)$ and $\text{Alt}(N)$ in reality have diameter polynomial in N (that is, polylogarithmic in their order, like the groups studied in the present paper).

A permutation group which provides a fascinating example intermediate between $\text{Sym}(N)$ and the groups considered in Theorems 1.1 and 1.3, is the Sylow p -subgroup $W_n = \text{Syl}_p(\text{Sym}(p^n))$ of $\text{Sym}(p^n)$. W_n was studied by Kaloujnine [22], who showed that it is isomorphic to the n -fold iterated regular wreath product $C_p \wr C_p \wr \cdots \wr C_p$, which acts naturally on the (first n levels of the) p -ary rooted tree. Although the inverse limit $\Gamma = \varprojlim_n W_n$ of the W_n is a regular branch pro- p group, it appears to be resistant to the Solovay-Kitaev procedure (for instance, Γ is not finitely generated as a topological group; by contrast, every version of the profinite Solovay-Kitaev procedure known to the author proves finite generation as a byproduct). It is however likely that the methods of this paper are applicable to other residually nilpotent branch groups. This should be a topic of further study.

The paper is structured as follows. Section 2 is devoted to preliminary material: in subsection 2.1 we lay out the consequences of the standard commutator identities upon which the profinite Solovay-Kitaev procedure is based, and illustrate, by means of an example, their relevance to diameter bounds. In subsection 2.2 we recall some basic material on group actions on regular rooted trees and the class of (regular) branch groups, define the Grigorchuk group \mathfrak{G} and the Gupta-Sidki p -groups $\Gamma_{(p)}$ and give some basic properties. In Section 3 we prove Theorem 1.2, and deduce Theorem 1.1. These proofs will be based in part upon prior results on the structure of the lower central series of \mathfrak{G} (taken from [3, 5]), which we state there. In Section 4 we prove Theorem 1.4 and deduce Theorem 1.3. In Section 5 we recall some results from [3] on the structure

of the lower central series of $\Gamma_{(3)}$ and using them, deduce Theorem 1.5 from Theorem 1.4. In Section 6 we comment on implications of our diameter bounds for spectral gap and mixing times of random walks on Cayley graphs. Finally in Section 7 we make some remarks on the relationship between the growth of an infinite group and the diameters of its finite quotients.

2 Preliminaries

2.1 Commutator Relations

The proofs of Theorems 1.2 and 1.4 (from which our other results are deduced) will be via a variant of the profinite Solovay-Kitaev Procedure developed in [8], to which we refer the reader for further background. Given a group G and a descending sequence of finite-index normal subgroups, the Procedure provides an approach to proving upper bounds on the diameters of the corresponding finite quotient groups, and relies on two key ingredients, both concerning the behaviour of commutator words within the group. The first ingredient encapsulates the intuitive notion that, given two specified elements $g, h \in G$ and two “good approximations” $\tilde{g}, \tilde{h} \in G$, the commutator $[\tilde{g}, \tilde{h}]$ of the approximations is a good approximation to the commutator $[g, h]$ of the original elements. What is perhaps not so intuitive is that in many situations, $[\tilde{g}, \tilde{h}]$ approximates $[g, h]$ much more closely than \tilde{g} and \tilde{h} did g and h . This idea is made precise in the following Lemma, which is an immediate consequence of standard commutator identities.

Lemma 2.1. *Let $G_1, G_2, H_1, H_2 \triangleleft G$, with $G_1 \geq G_2$, $H_1 \geq H_2$. For all $g_i \in G_i$, $h_i \in H_i$ ($i = 1, 2$),*

$$[g_1 g_2, h_1 h_2] \equiv [g_1, h_1] \pmod{[G_1, H_2][G_2, H_1]}.$$

Proof. We compute directly:

$$[g_1 g_2, h_1 h_2] = [g_1, h_2][g_1, h_1][[g_1, h_1], h_2][[g_1, h_1 h_2], g_2][g_2, h_1 h_2]$$

and all terms other than $[g_1, h_1]$ lie in $[G_1, H_2][G_2, H_1]$. \square

A particularly useful special case of this lemma is the following.

Corollary 2.2. *Let $(K_i)_{i=1}^\infty$ be a descending sequence of normal subgroups of G . Suppose that, for all $m, n \in \mathbb{N}$, $[K_m, K_n] \subseteq K_{m+n}$. Let $m_1, m_2, n_1, n_2 \in \mathbb{N}$, with $m_1 \leq m_2$, $n_1 \leq n_2$, and let $g_i \in K_{m_i}$, $h_i \in K_{n_i}$ ($i = 1, 2$). Then:*

$$[g_1 g_2, h_1 h_2] \equiv [g_1, h_1] \pmod{K_{\min(m_1+n_2, m_2+n_1)}}.$$

Proof. Set $G_i = K_{m_i}$ and $H_i = K_{n_i}$ in Lemma 2.1, for $i = 1, 2$. \square

We will use Corollary 2.2 in the proof of Theorem 1.2. It is a classical fact that the lower central series satisfies the required hypothesis.

Lemma 2.3. *Let G be any group. Then for all $m, n \in \mathbb{N}$,*

$$[\gamma_n(G), \gamma_m(G)] \subseteq \gamma_{m+n}(G).$$

Lemma 2.3 will be used extensively and without further comment in the sequel, and in particular throughout Section 3.

What is the relevance of the preceding discussion to diameters? As we intimated in the introduction, given a group Γ and a descending sequence $(\Gamma_i)_i$, the Solovay-Kitaev procedure requires the existence of an approximation to elements of a deeper term Γ_j in the sequence, by commutators $[g, h]$ of elements g, h lying in a higher term Γ_i . We may assume by induction that we have approximations \tilde{g}, \tilde{h} to g, h up to an error lying in the intermediate term Γ_k , where \tilde{g} and \tilde{h} are short words in a generating set. Substituting \tilde{g} and \tilde{h} into our commutator expression, we obtain approximations $[\tilde{g}, \tilde{h}]$ to elements of Γ_j by short words. Lemma 2.1 gives us some control over the fidelity of these approximations. Let us illustrate this with an example.

Example 2.4. Let Γ be a group, let $\Gamma \geq \Gamma_1 \geq \Gamma_2 \geq \Gamma_3$ be finite-index normal subgroups of Γ , and suppose:

$$[\Gamma_1, \Gamma_2] \leq \Gamma_3. \quad (1)$$

Second, suppose that for any $k \in \Gamma_2$, there exist $g, h \in \Gamma_1$ such that:

$$[g, h]k^{-1} \in \Gamma_3. \quad (2)$$

Now let $S \subseteq \Gamma$, and suppose the image of S in Γ/Γ_3 is a generating set. Let $d = \text{diam}(\Gamma/\Gamma_2, S)$, so that:

$$\Gamma = \Gamma_2 B_S(d) \quad (3)$$

(here we abuse notation slightly, denoting also by “ S ” the image of S modulo any of the Γ_i).

We wish to bound $\text{diam}(\Gamma/\Gamma_3, S)$. That is, given $k \in \Gamma$ we seek a short word w in S such that $kw^{-1} \in \Gamma_3$. First suppose that $k \in \Gamma_2$. Let $g, h \in \Gamma_1$ be as in (2). By (3), there exist $\tilde{g}, \tilde{h} \in B_S(d)$ such that $g^{-1}\tilde{g}, h^{-1}\tilde{h} \in \Gamma_2$. Setting $G_1 = H_1 = \Gamma_1$, $G_2 = H_2 = \Gamma_2$, $g_1 = g$, $h_1 = h$, $g_2 = g^{-1}\tilde{g}$, $h_2 = h^{-1}\tilde{h}$ in Lemma 2.1, and applying (1),

$$[\tilde{g}, \tilde{h}] \equiv [g, h] \equiv k \pmod{\Gamma_3}.$$

Since $[\tilde{g}, \tilde{h}] \in B_S(4d)$, we have $\Gamma_2 \subseteq \Gamma_3 B_S(4d)$. Applying (3) again, $\Gamma = \Gamma_3 B_S(5d)$, so $\text{diam}(\Gamma/\Gamma_3, S) \leq 5d$.

Although the hypotheses (1) and (2) may seem somewhat artificial in the context of this abstract example, in practice many groups satisfy these conditions, or variants thereof. Modifications to the method of Example 2.4 are however sometimes necessary, or desirable. For instance the elements g and h in Example 2.4 were taken to lie at the same level of the subgroup chain, whereas for many groups, including the branch groups studied in this paper, good commutator expressions will involve elements lying at different levels.

2.2 Groups Acting on Regular Rooted Trees

Let us set up some basic notation and definitions concerning the class of groups to be studied in the sequel. All of the material here (and much, much more besides) is covered in [6] and [10] Chapter VIII.

Let \mathcal{A} be a finite set, and let \mathcal{A}^* be the set of formal positive words on the alphabet \mathcal{A} . We may partially order \mathcal{A}^* via the *prefix relation* \leq , where for $u, v \in \mathcal{A}^*$, $u \leq v$ iff there exists $w \in \mathcal{A}^*$ such that $v = uw$.

Geometrically, we may regard \mathcal{A}^* as the set of vertices of a regular rooted tree $\mathcal{T}_{\mathcal{A}}$: the root vertex is identified with the empty word, and every vertex v is joined by an edge to its $|\mathcal{A}|$ children va , $a \in \mathcal{A}$. Under this identification, the set \mathcal{A}^n of words of length n is precisely the sphere of radius n in $\mathcal{T}_{\mathcal{A}}$ about the root vertex, known as the *n th level set*.

The *automorphism group* $\text{Aut}(\mathcal{T}_{\mathcal{A}})$ is the set of permutations of \mathcal{A}^* preserving the prefix relation. Geometrically this is just the group of graph automorphisms of the tree $\mathcal{T}_{\mathcal{A}}$. For the valence of every vertex of $\mathcal{T}_{\mathcal{A}}$ is $|\mathcal{A}| + 1$ except for the root vertex (which has valence $|\mathcal{A}|$), so every automorphism of $\mathcal{T}_{\mathcal{A}}$ fixes the root vertex, and hence preserves the level sets. The kernel of the action of $\text{Aut}(\mathcal{T}_{\mathcal{A}})$ on the n th level set \mathcal{A}^n will be called the *n th level stabiliser* and denoted $\text{Stab}(n)$; it is naturally isomorphic to $\text{Aut}(\mathcal{T}_{\mathcal{A}})^{(\times |\mathcal{A}|^n)}$. If $\Gamma \leq \text{Aut}(\mathcal{T}_{\mathcal{A}})$ we write $\text{Stab}_{\Gamma}(n)$ for $\Gamma \cap \text{Stab}(n)$, though in general we cannot say more about the structure of $\text{Stab}_{\Gamma}(n)$ than that $\text{Stab}_{\Gamma}(n)$ is isomorphic to a subgroup of $\text{Aut}(\mathcal{T}_{\mathcal{A}})^{(\times |\mathcal{A}|^n)}$.

For any $\phi \in \text{Aut}(\mathcal{T}_{\mathcal{A}})$, there exists a unique $\sigma_{\phi} \in \text{Sym}(\mathcal{A})$ such that for any $x \in \mathcal{A}$, there exists a unique $\phi_x \in \text{Aut}(\mathcal{T}_{\mathcal{A}})$ such that:

$$\phi(xw) = \sigma_{\phi}(x)\phi_x(w), \text{ for all } w \in \mathcal{A}^*.$$

The induced map $\psi : \phi \mapsto (\phi_x)_{x \in \mathcal{A}} \cdot \sigma_{\phi}$ gives an isomorphism $\text{Aut}(\mathcal{T}_{\mathcal{A}}) \rightarrow \text{Aut}(\mathcal{T}_{\mathcal{A}}) \wr \text{Sym}(\mathcal{A})$. Note that the level stabilisers may be described recursively by $\text{Stab}(0) = \text{Aut}(\mathcal{T}_{\mathcal{A}})$ and $\text{Stab}(n+1) = \psi^{-1}(\text{Stab}(n)^{(\times |\mathcal{A}|)})$ for $n \in \mathbb{N}$.

Of particular interest among the subgroups of $\text{Aut}(\mathcal{T}_{\mathcal{A}})$ are those whose action on \mathcal{A}^* is *branch*. Our characterization of such groups is based on that appearing in [3].

Definition 2.5. Let $\Gamma \leq \text{Aut}(\mathcal{T}_{\mathcal{A}})$. Γ is (regular) *branch* if:

- (i) The action of Γ on \mathcal{A} is transitive;
- (ii) $\psi(\text{Stab}_{\Gamma}(1)) \leq \Gamma^{(\times |\mathcal{A}|)}$;
- (iii) Γ has a finite-index subgroup K such that $K^{(\times |\mathcal{A}|)} \leq \psi(K)$.

We will simply say that a group Γ *branches over K* when the alphabet \mathcal{A} and the action of Γ on \mathcal{A}^* is clear.

For the sake of uncluttered notation we will allow ourselves to suppress the map ψ from expressions and identify subgroups of Γ with their image under ψ , so for instance we may (abuse notation somewhat and) speak of $K^{(\times |\mathcal{A}|)}$ as a subgroup of K ; $\text{Stab}_{\Gamma}(n)$ as a subgroup of $\Gamma^{(\times |\mathcal{A}|^n)}$ and so on.

In truth, branch groups form a much broader class than those groups covered by Definition 2.5 (see for instance [6]), but this more restricted setting will be most convenient for our purposes.

We now define the specific branch groups which are the subject of Theorems 1.1-1.5.

2.2.1 The Grigorchuk 2-Group

Let $\mathcal{A} = \{0, 1\}$ and write $\mathcal{T}_{\mathcal{A}} = \mathcal{T}_2$. The *Grigorchuk 2-group* (sometimes known as the *first Grigorchuk group*) is the subgroup \mathfrak{G} of $\text{Aut}(\mathcal{T}_2)$ generated by the four automorphisms a, b, c, d , defined by:

$$\begin{aligned} a(0w) &= 1w; a(1w) = 0w; \\ b(0w) &= 0a(w); b(1w) = 1c(w); \\ c(0w) &= 0a(w); c(1w) = 1d(w); \\ d(0w) &= 0w; d(1w) = 1b(w). \end{aligned}$$

In other words, a swaps the subtrees rooted at 0 and 1, while $b, c, d \in \text{Stab}_{\mathfrak{G}}(1)$ are defined recursively via:

$$b = (a, c), c = (a, d), d = (1, b).$$

An easy induction on the levels shows that:

$$a^2 = b^2 = c^2 = d^2 = 1, bc = cb = d, cd = dc = b, bd = db = c. \quad (4)$$

Let $x = abab \in \mathfrak{G}$ and let $K = \langle x \rangle^{\mathfrak{G}} \triangleleft \mathfrak{G}$. We will require the following basic facts in the sequel.

Proposition 2.6 ([10] Chapter VIII). *Let \mathfrak{G} , K , x be as above.*

- (i) \mathfrak{G} branches over K ;
- (ii) $\mathfrak{G}/K \cong D_8 \times C_2$;
- (iii) $K/K^{(\times 2)} \cong C_4$, generated by x .

Lemma 2.7. $K^{(\times 2^m)} \triangleleft \mathfrak{G}$ for all $m \geq 0$.

Proof. We proceed by induction on m , the base case $m = 0$ holding by definition.

For $m \geq 1$, we verify that $K^{(\times 2^m)}$ is preserved under conjugation by the generators a, b, c, d . This is the case for a , which simply permutes the factors in the direct product.

b, c and d preserve the decomposition $K^{(\times 2^m)} = (K^{(\times 2^{m-1})})^{(\times 2)}$, and act on each $K^{(\times 2^{m-1})}$ -factor as a, b, c or d . The result follows by induction. \square

Lemma 2.8. For all $m \geq 0$, $K^{(\times 2^m)} \leq \text{Stab}_{\mathfrak{G}}(m+1)$.

Proof. Note that $x = abab \in \text{Stab}_{\mathfrak{G}}(1)$. The result is now immediate from Proposition 2.6 (iii). \square

Lemma 2.9. *For all $n \geq 1$,*

$$|\mathfrak{G} : \text{Stab}_{\mathfrak{G}}(n)| \geq 2^{2^{n-1}+1}.$$

Proof. It suffices to prove that $|\text{Stab}_{\mathfrak{G}}(m) : \text{Stab}_{\mathfrak{G}}(m+1)| \geq 2^{2^{m-1}}$ for all $m \in \mathbb{N}$. $x = abab \equiv (a, a) \pmod{\text{Stab}_{\text{Aut}(\mathcal{T}_2)}(2)}$, so as \underline{e} ranges over $\{0, 1\}^{2^m}$, the elements $(x^{\underline{e}_i})_{i=1}^{2^m}$ are all distinct modulo $\text{Stab}_{\mathfrak{G}}(m+1)$, and lie in $\text{Stab}_{\mathfrak{G}}(m)$ by Lemma 2.8. \square

2.2.2 The Gupta-Sidki p -Groups

Fix an odd prime p , let $\mathcal{A} = \{0, 1, \dots, p-1\}$ and write $\mathcal{T}_{\mathcal{A}} = \mathcal{T}_p$. The *Gupta-Sidki p -group* is the subgroup $\Gamma_{(p)}$ of $\text{Aut}(\mathcal{T}_p)$ generated by the two automorphisms a, b , where a is defined by:

$$a(iw) = (i+1)w \text{ for } 0 \leq i \leq p-2; a((p-1)w) = 0w$$

(so that a cyclically permutes the level-1 subtrees) and $b \in \text{Stab}_{\text{Aut}(\mathcal{T}_p)}(1)$ is defined recursively via:

$$b = (a, a^{-1}, 1, \dots, 1, b)$$

so that both a and b have order p . Note that a reduced word in a and b corresponds to an element of $\text{Stab}_{\Gamma_{(p)}}(1)$ iff the number of occurrences of a (counted with signs) is congruent to 0 modulo p . Thus $\text{Stab}_{\Gamma_{(p)}}(1) = \langle b \rangle^{\Gamma_{(p)}}$.

Let $K = [\Gamma_{(p)}, \Gamma_{(p)}]$ be the derived subgroup of $\Gamma_{(p)}$. Note that $K \leq \text{Stab}_{\Gamma_{(p)}}(1)$. Let $x_1 = [a, b] \in K$, and for $1 \leq i \leq p-2$, define $x_{i+1} = [a, x_i] \in K$. The following computations were made in [15], generalising results from [27] for $\Gamma_{(3)}$.

Proposition 2.10 ([15] Proposition 2.2). *Let $B = \text{Stab}_{\Gamma_{(p)}}(1)$. Then:*

- (i) $\Gamma_{(p)}/K \cong C_p \times C_p$, with basis Ka, Kb ;
- (ii) $B' = K^{(\times p)}$;
- (iii) $\Gamma_{(p)}/B' \cong C_p \wr C_p$.

From this we have an explicit description of the branch structure of $\Gamma_{(p)}$.

Corollary 2.11. (i) $\Gamma_{(p)}$ branches over K ;

- (ii) $K/K^{(\times p)} \cong C_p^{(\times(p-1))}$, with basis x_1, \dots, x_{p-1} .

Proof. (i) From Proposition 2.10 (ii),

$$K^{(\times p)} = B' \leq \Gamma'_{(p)} = K.$$

- (ii) We defer this to Section 4, where we introduce additional notation which will be convenient to use in the proof. \square

The following observation will be key to the deduction of Theorem 1.3 from Theorem 1.4.

Lemma 2.12. *For all $m \in \mathbb{N}$, $K^{(\times p^m)} \leq \text{Stab}_{\Gamma_{(p)}}(m+1)$.*

Proof. It clearly suffices to check that $K \leq \text{Stab}_{\Gamma_{(p)}}(1)$. This is so because $K = [\Gamma_{(p)}, \Gamma_{(p)}]$ and $\Gamma_{(p)}/\text{Stab}_{\Gamma_{(p)}}(1)$ is abelian. \square

3 Proofs for Grigorchuk's Group

In this Section we prove Theorem 1.2 and deduce Theorem 1.1. Before embarking on the proof of our diameter bounds, we marshal some facts about the lower central series of \mathfrak{G} .

Recall that for any group G , the degree $\deg(g)$ of $g \in G$ is given by $g \in \gamma_{\deg(g)}(G) \setminus \gamma_{\deg(g)+1}(G)$, with the convention that all $g \in \bigcap_{n=1}^{\infty} \gamma_n(G)$ have degree ∞ (though in \mathfrak{G} , the latter situation never arises for $g \neq 1$: \mathfrak{G} is a residually finite 2-group, so is residually 2-finite, and in particular, $\bigcap_{n=1}^{\infty} \gamma_n(\mathfrak{G}) = \{1\}$).

For any $g \in K$, write $\mathbf{0}(g) = (g, 1)$, $\mathbf{1}(g) = (g, g^{-1}) \in K^{(\times 2)}$.

Theorem 3.1 ([3, 5, 26]). *Let $X_1, \dots, X_n \in \{\mathbf{0}, \mathbf{1}\}$. Then:*

$$\begin{aligned} \deg X_1 \cdots X_n(x) &= 1 + \sum_{i=1}^n X_i 2^{i-1} + 2^n; \\ \deg X_1 \cdots X_n(x^2) &= 1 + \sum_{i=1}^n X_i 2^{i-1} + 2^{n+1}. \end{aligned}$$

Theorem 3.2 ([3, 5, 26]). *For all $m \geq 2$,*

$$|\mathfrak{G} : \gamma_{2^m+1}(\mathfrak{G})| = 2^{2^{m-1}3+2}.$$

To avoid cluttered notation, in this section we may write $\gamma_n = \gamma_n(\mathfrak{G})$. The specific consequences of Theorem 3.1 to be used in the proof of Theorem 1.2 are as follows.

Corollary 3.3. $\gamma_{2^m+2^{m-1}+1} \leq K^{(\times 2^m)} \leq \gamma_{2^m+1}$.

To be more precise, we have the following estimates.

Corollary 3.4. *For all $\underline{\delta} \in (\{0, 1\}^{2^{n-2}}) \setminus \{\underline{0}\}$,*

- (i) $2^n + 1 \leq \deg \left(((x, 1, 1, 1)^{\delta_i})_{i=1}^{2^{n-2}} \right) \leq 2^n + 2^{n-2};$
- (ii) $2^n + 2^{n-2} + 1 \leq \deg \left(((x, 1, x, 1)^{\delta_i})_{i=1}^{2^{n-2}} \right) \leq 2^n + 2^{n-1};$
- (iii) $2^n + 2^{n-1} + 1 \leq \deg \left(((x, x, 1, 1)^{\delta_i})_{i=1}^{2^{n-2}} \right) \leq 2^n + 3 \cdot 2^{n-2};$
- (iv) $2^n + 3 \cdot 2^{n-2} + 1 \leq \deg \left(((x, x, x, x)^{\delta_i})_{i=1}^{2^{n-2}} \right) \leq 2^{n+1};$

- (v) Specifically, $\deg((x, x, x, x)_{i=1}^{2^{n-2}}) = 2^{n+1}$;
- (vi) $2^n + 1 \leq \deg((x^2, 1)^{\delta_i})_{i=1}^{2^{n-2}} \leq 2^n + 2^{n-2}$;
- (vii) $2^n + 2^{n-2} + 1 \leq \deg((x^2, x^2)^{\delta_i})_{i=1}^{2^{n-2}} \leq 2^n + 2^{n-1}$;
- (viii) Specifically, $\deg((x^2, x^2)_{i=1}^{2^{n-2}}) = 2^n + 2^{n-1}$.

We now produce our commutator approximations to elements lying sufficiently deep in $(\gamma_n)_n$. The following identities are verified by direct computation.

Lemma 3.5. (i) $[x, (x, 1)] = (x^{-1}, 1, 1, 1)$.

(ii) $[x, (x, x)] = (x^{-1}, 1, 1, (1, x^{-1})x)$.

(iii) $[x^2, (x, 1)] = (x^{-1}, x, 1, 1)$.

(iv) $[x^2, (x, x^{-1})] = (x^{-1}, x, (x^{-1}, 1)x^{-1}, (1, x^{-1})x)$.

Proof. We prove (i) and leave the verifications of the other identities, which are similar, as an exercise. First note that $x = abab = (ca, ac)$. Thus:

$$[x, (x, 1)] = ([ca, x], 1)$$

Now, using (4) we have:

$$\begin{aligned} [ca, x] &= ac(ac, ca)ca(ca, ac) \\ &= a(ca, bad)a(ca, ac) \\ &= (bad, ca)(ca, ac) \\ &= (baba, 1) \\ &= (x^{-1}, 1) \end{aligned}$$

as desired. □

Lemma 3.5 yields good approximations to elements of $K^{(\times 4)}$, modulo $K^{(\times 8)}$, by commutators of elements of K and $K^{(\times 2)}$. For deeper subgroups $K^{(\times 2^n)}$, we express elements as vectors of elements in $K^{(\times 4)}$, and produce a commutator approximation by applying Lemma 3.5 to each term of the vector. The expressions we obtain can be related to the lower central series by using Theorem 3.1 and Corollaries 3.3 and 3.4 to estimate the degrees of the elements occurring. In summary we have the following Proposition.

Proposition 3.6. Let $\mathfrak{c} : \mathfrak{G} \times \mathfrak{G} \rightarrow \mathfrak{G}$ be given by $\mathfrak{c}(g, h) = [g, h]$. Let $m \geq 2$. Then:

- (i) The restriction of \mathfrak{c} to $\gamma_{2^{m-1}} \times \gamma_{2^{m-1}}$ descends to a well-defined map:

$$\bar{\mathfrak{c}}_m : (\gamma_{2^{m-1}}/\gamma_{2^m+1})^{(\times 2)} \rightarrow \mathfrak{G}/\gamma_{2^m+2^{m-1}+1}$$

whose image contains $K^{(\times 2^m)}/\gamma_{2^m+2^{m-1}+1}$.

(ii) The restriction of \mathfrak{c} to $\gamma_{2^{m-1}+2^{m-2}} \times \gamma_{2^{m-1}+2^{m-2}}$ descends to a well-defined map:

$$\bar{\mathfrak{c}}_m : (\gamma_{2^{m-1}+2^{m-2}}/\gamma_{2^m+2^{m-1}+1})^{(\times 2)} \rightarrow \mathfrak{G}/\gamma_{2^{m+1}+1}$$

whose image contains $\gamma_{2^m+2^{m-1}+1}/\gamma_{2^{m+1}+1}$.

Proof. The well-definedness of $\bar{\mathfrak{c}}_m$ and $\bar{\mathfrak{c}}_m$ is an immediate consequence of Corollary 2.2. It therefore suffices to check the images of the maps contain the specified subgroups.

For (i), note that by Corollary 3.4, every element of $K^{(\times 2^m)}/\gamma_{2^m+2^{m-1}+1}$ is represented by $((x^{\delta_i+\epsilon_i}, 1, x^{\epsilon_i}, 1))_{i=1}^{2^{m-2}}$ for some $\delta_i, \epsilon_i \in \{0, 1\}$. By Lemma 3.5 (i) and (ii),

$$\begin{aligned} [x, (x, 1)] \cdot (x, 1, 1, 1)^{-1} &= (x^2, 1, 1, 1)^{-1} \\ [x, (x, x)] \cdot (x, 1, x, 1)^{-1} &\equiv (1, 1, x, x)(x^2, 1, x^2, 1)^{-1} \pmod{K^{(\times 2^3)}}. \end{aligned}$$

By Corollary 3.4 (iii) and (iv), for any $(\beta_i)_{i=1}^{2^{m-2}} \in \{0, 1\}^{2^{m-2}}$,

$$((1, 1, x, x)^{\beta_i})_{i=1}^{2^{m-2}} \in \gamma_{2^m+2^{m-1}+1},$$

by Corollary 3.4 (vi), for any $(\beta_i)_{i=1}^{2^{m-2}} \in \{0, 1\}^{2^{m-2}}$,

$$((x^2, 1, 1, 1)^{\beta_i})_{i=1}^{2^{m-2}}, ((x^2, 1, x^2, 1)^{\beta_i})_{i=1}^{2^{m-2}} \in \gamma_{2^m+2^{m-1}+1},$$

and from Corollary 3.3, $K^{(\times 2^{m+1})} \subseteq \gamma_{2^{m+1}+1}$. Hence, for any $(\delta_i)_{i=1}^{2^{m-2}}, (\epsilon_i)_{i=1}^{2^{m-2}} \in \{0, 1\}^{2^{m-2}}$,

$$\begin{aligned} ((x^{\delta_i+\epsilon_i}, 1, x^{\epsilon_i}, 1))_{i=1}^{2^{m-2}} &\equiv [(x)_{i=1}^{2^{m-2}}, ((x, 1)^{\delta_i})_{i=1}^{2^{m-2}}] \\ &\quad \cdot [(x)_{i=1}^{2^{m-2}}, ((x, x)^{\epsilon_i})_{i=1}^{2^{m-2}}] \pmod{\gamma_{2^m+2^{m-1}+1}}. \end{aligned}$$

Using Corollary 3.4 to estimate the degrees of $(x)_{i=1}^{2^{m-2}}, ((x, 1)^{\delta_i})_{i=1}^{2^{m-2}}$ and $((x, x)^{\epsilon_i})_{i=1}^{2^{m-2}}$, and by the standard identity $[a, bc] = [a, c][a, b][[a, b], c]$, we deduce:

$$((x^{\delta_i+\epsilon_i}, 1, x^{\epsilon_i}, 1))_{i=1}^{2^{m-2}} \equiv [(x)_{i=1}^{2^{m-2}}, ((x^{\delta_i+\epsilon_i}, x^{\epsilon_i}))_{i=1}^{2^{m-2}}] \pmod{\gamma_{2^m+2^{m-1}+1}}. \quad (5)$$

For (ii), we see similarly by Corollary 3.4 that every element of $\gamma_{2^m+2^{m-1}+1}/\gamma_{2^{m+1}+1}$ is represented by $((x^{\delta_i+\epsilon_i}, x^{\delta_i+\epsilon_i}, x^{\epsilon_i}, x^{\epsilon_i}))_{i=1}^{2^{m-2}}$ for some $\delta_i, \epsilon_i \in \{0, 1\}$. By Lemma 3.5 (iii) and (iv),

$$\begin{aligned} [x^2, (x, 1)] \cdot (x, x, 1, 1)^{-1} &= (x^2, 1, 1, 1)^{-1} \\ [x^2, (x, x^{-1})] \cdot (x, x, x, x)^{-1} &\equiv (x^2, 1, x^2, 1)^{-1} \pmod{K^{(\times 2^3)}}. \end{aligned}$$

Using Corollary 3.3 and Corollary 3.4 (vi) once again as in (i), we have that for any $(\delta_i)_{i=1}^{2^{m-2}}, (\epsilon_i)_{i=1}^{2^{m-2}} \in \{0, 1\}^{2^{m-2}}$,

$$\begin{aligned} ((x^{\delta_i+\epsilon_i}, x^{\delta_i+\epsilon_i}, x^{\epsilon_i}, x^{\epsilon_i}))_{i=1}^{2^{m-2}} &\equiv [(x^2)_{i=1}^{2^{m-2}}, ((x, 1)^{\delta_i})_{i=1}^{2^{m-2}}] \\ &\quad \cdot [(x^2)_{i=1}^{2^{m-2}}, ((x, x^{-1})^{\epsilon_i})_{i=1}^{2^{m-2}}] \pmod{\gamma_{2^m+2^{m-1}+1}}. \end{aligned}$$

As before, we apply the commutator identity for products, using the estimate of the degrees of $(x^2)_{i=1}^{2^{m-2}}, ((x, 1)^{\delta_i})_{i=1}^{2^{m-2}}$ and $((x, x^{-1})^{\epsilon_i})_{i=1}^{2^{m-2}}$ from Corollary 3.4, and deduce:

$$((x^{\delta_i+\epsilon_i}, x^{\delta_i+\epsilon_i}, x^{\epsilon_i}, x^{\epsilon_i}))_{i=1}^{2^{m-2}} \equiv [(x^2)_{i=1}^{2^{m-2}}, ((x^{\delta_i+\epsilon_i}, x^{-\epsilon_i}))_{i=1}^{2^{m-2}}] \pmod{\gamma_{2^{m+1}+1}}. \quad (6)$$

□

Using Proposition 3.6 we may approximate any tuple consisting of x s and 1s by a commutator. To express arbitrary tuples in $K^{(\times 2^{m-1})}/K^{(\times 2^m)}$ we must also find approximations for tuples consisting of x^2 s and 1s. Here we will diverge slightly from our overall strategy of approximating elements by commutators, since it appears far more natural to express such tuples as squares.

Proposition 3.7. *The squaring map $\mathfrak{s} : g \mapsto g^2$ on \mathfrak{G} induces a surjection:*

$$\overline{\mathfrak{s}}_m : K^{(\times 2^{m-1})}/\gamma_{2^m+1} \rightarrow \gamma_{2^m+1}/K^{(\times 2^m)}$$

for all $m \geq 1$.

Proof. It is immediate from Lemma 2.6 and Corollaries 3.3 and 3.4 that $K^{(\times 2^{m-1})}/\gamma_{2^m+1}$ and $\gamma_{2^m+1}/K^{(\times 2^m)}$ are elementary abelian 2-groups, each element of the latter being represented by a vector:

$$(x^{2\delta_i})_{i=1}^{2^{m-1}} = \mathfrak{s}((x^{\delta_i})_{i=1}^{2^{m-1}}),$$

as δ ranges over $\{0, 1\}^{2^{m-1}}$. Since $(x^{\delta_i})_{i=1}^{2^{m-1}} \in K^{(\times 2^{m-1})}$, \mathfrak{s} induces a surjection $K^{(\times 2^{m-1})} \rightarrow \gamma_{2^m+1}/K^{(\times 2^m)}$.

Now let $a \in K^{(\times 2^{m-1})}$, $b \in \gamma_{2^m+1}$. We have:

$$\mathfrak{s}(ab) = \mathfrak{s}(a)[a, b^{-1}]\mathfrak{s}(b).$$

But $\mathfrak{s}(\gamma_{2^m+1}) \subseteq K^{(\times 2^m)}$ (as noted above) and:

$$[K^{(\times 2^{m-1})}, \gamma_{2^m+1}] \subseteq K^{(\times 2^m)}$$

by Corollary 3.3 (applied to both $K^{(\times 2^{m-1})}$ and $K^{(\times 2^m)}$). Thus $\overline{\mathfrak{s}}_m$ is indeed well-defined. □

We now come to the heart of the proof of our diameter bound: using Propositions 3.6 and 3.7, we show that if a symmetric subset $X \subseteq \mathfrak{G}$ contains an approximation to every element of \mathfrak{G} up to an error lying in γ_{2^m+1} , then every element of \mathfrak{G} is approximated, up to an error in the (much smaller) subgroup $\gamma_{2^{m+1}+1}$, by a short word in X .

Proposition 3.8. *Let $m \geq 2$ and let $X \subseteq \Gamma$ be a symmetric subset such that:*

$$X\gamma_{2^m+1} = \mathfrak{G}. \quad (7)$$

Then:

$$X^{35}\gamma_{2^{m+1}+1} = \mathfrak{G}. \quad (8)$$

Proof. First, note that Proposition 3.7 immediately implies:

$$X^2K^{(\times 2^m)} \supseteq \gamma_{2^m+1}. \quad (9)$$

Second, we combine (7) and Proposition 3.6 (ii) with Corollary 2.2 to conclude:

$$K^{(\times 2^m)} \subseteq X^4\gamma_{2^m+2^{m-1}+1}. \quad (10)$$

Taking stock of what we have thus far, (7), (9) and (10) combine to give:

$$X^7\gamma_{2^m+2^{m-1}+1} = \mathfrak{G}. \quad (11)$$

Finally, we combine (11) and (6) with Corollary 2.2 and conclude:

$$\gamma_{2^m+2^{m-1}+1} \subseteq X^{28}\gamma_{2^{m+1}+1}. \quad (12)$$

The required conclusion (8) is now immediate from (11) and (12). \square

Proof of Theorem 1.2. Let $S \subseteq \mathfrak{G}/\gamma_n(\mathfrak{G})$ be a generating set. If $n \leq 5$ then:

$$\text{diam}(\mathfrak{G}/\gamma_n(\mathfrak{G}), S) \leq |\mathfrak{G} : \gamma_n(\mathfrak{G})|$$

is bounded by an absolute constant \tilde{C} . Otherwise let $\tilde{S} \subseteq \mathfrak{G}$ be any subset whose image in $\mathfrak{G}/\gamma_n(\mathfrak{G})$ is S , and let $m \in \mathbb{N}$ be such that $2^{m-1} + 1 < n \leq 2^m + 1$. Then $B_{\tilde{S}}(\tilde{C})\gamma_5(\mathfrak{G}) = \mathfrak{G}$, and by repeated application of Proposition 3.8,

$$B_{\tilde{S}}(35^{m-2}\tilde{C})\gamma_{2^m+1}(\mathfrak{G}) = \mathfrak{G}$$

so that $\text{diam}(\mathfrak{G}/\gamma_n(\mathfrak{G}), S) \leq 35^{m-2}\tilde{C} \ll n^{\log(35)/\log(2)}$.

The result now follows from Theorem 3.2. \square

Remark 3.9. *Note that the above proof facilitates straightforward computation of the implied constant from the statement of Theorem 1.2.*

Proof of Theorem 1.1. By Lemma 2.8 and Corollary 3.3,

$$\gamma_{2^{n+1}+1}(\mathfrak{G}) \leq \text{Stab}_{\mathfrak{G}}(n+1)$$

so:

$$\text{diam}(\mathfrak{G}/\text{Stab}_{\mathfrak{G}}(n+1)) \leq \text{diam}(\mathfrak{G}/\gamma_{2^{n+1}+1}(\mathfrak{G})).$$

The result is now immediate from Theorem 1.2, Lemma 2.9 and Theorem 3.2. \square

Remark 3.10. It is very likely that detailed knowledge of the lower central series of \mathfrak{G} is not required to prove Theorem 1.1. Rather, one could give a direct proof of a diameter bound for $\mathfrak{G}/K^{(\times 2^n)}$ and deduce Theorem 1.1 from this, much as we do for the Gupta-Sidki groups in the following Section. The reason for organizing the proof of Theorem 1.1 as it appears here is historical. Theorem 1.2 was the first of our results to be proved, followed by a direct proof of Theorem 1.5, using the results of [3] on the lower central series of $\Gamma_{(3)}$. The case $p = 3$ of Theorem 1.3 was then deduced from Theorem 1.5 (much as Theorem 1.1 follows from Theorem 1.2). The (arguably more natural) proof of Theorem 1.3 from Theorem 1.4 was a later response to the need to avoid assuming knowledge of the lower central series of $\Gamma_{(p)}$ in proving Theorem 1.3 for higher p (to the author's knowledge, the lower central series of $\Gamma_{(p)}$ has not been computed for $p \geq 5$).

Remark 3.11. Proposition 3.7 hints at the possibility of an alternative approach to proving diameter bounds for sequences of groups, following the same broad lines as the Solovay-Kitaev procedure employed here and in [8], but using power-words instead of commutator words. This will be explored further elsewhere [9].

4 Proofs for the Gupta-Sidki Groups

In this section we prove Theorem 1.4 and, from it, Theorem 1.3. For the remainder of the section we shall write Γ for $\Gamma_{(p)}$. Any assumptions on the prime p in what follows will be made explicit in the appropriate place.

The following notation will be useful in the sequel: for any $g \in \text{Aut}(\mathcal{T}_{\mathcal{A}})$, let $\mathbf{0}(g) = (g, 1, \dots, 1) \in \text{Aut}(\mathcal{T}_{\mathcal{A}})^{(\times p)}$, and for $0 \leq j \leq p-1$, let $(\mathbf{j} + \mathbf{1})(g) = [a, \mathbf{j}(g)]$. Hence for $0 \leq j \leq p-1$,

$$\mathbf{j}(g)_i = \begin{cases} g^{\alpha_{j,i}} & 1 \leq i \leq j+1 \\ 1 & \text{otherwise} \end{cases}, \text{ where } \alpha_{j,i} = (-1)^{i+1} \binom{j}{i-1}.$$

Note that $\alpha_{p-1,i} \equiv 1 \pmod p$ for all $1 \leq i \leq p$. It follows that:

$$(\mathbf{p} - \mathbf{1})(a) = (a, \dots, a), (\mathbf{p} - \mathbf{1})(b) = (b, \dots, b)$$

(since a, b have order p) and:

$$(\mathbf{p} - \mathbf{1})(g) \equiv (g, \dots, g) \pmod{K^{(\times p^2)}} \text{ for any } g \in K$$

(since $K \leq \text{Stab}_{\Gamma}(1)$ and Γ/K has exponent p , so too does $K/K^{(\times p)}$).

Note also that in this notation,

$$b = \mathbf{1}(a)\mathbf{0}(b)^{a^{-1}} = \mathbf{0}(b)^{a^{-1}}\mathbf{1}(a).$$

It follows from the definition of the elements x_i that for $1 \leq i \leq p-1$:

$$x_i((\mathbf{i} + \mathbf{1})(a)\mathbf{i}(b)^{a^{-1}})^{-1} \in K^{(\times p)} \quad (13)$$

and in particular:

$$x_{p-2}(a, \dots, a)^{-1} \in (\text{Stab}_\Gamma(1))^{(\times p)} \quad (14)$$

$$x_{p-1}(b, \dots, b)^{-1} \in K^{(\times p)} \quad (15)$$

These observations facilitate the completion of:

Proof of Corollary 2.11 (ii). Since $K \leq \text{Stab}_\Gamma(1)$, $K/K^{(\times p)}$ is an elementary abelian p -group. Moreover, from Proposition 2.10 (i) and (iii) we have $|K : K^{(\times p)}| = p^{p-1}$, so it suffices to check that the images of x_1, \dots, x_{p-1} in $K/K^{(\times p)}$ are linearly independent.

Embedding $\text{Stab}_\Gamma(1)$ into $\Gamma^{(\times p)}$ (via ψ), we have an induced embedding of $K/K^{(\times p)}$ into $(\Gamma/K)^{(\times p)}$. By Proposition 2.10 (i),

$$(\Gamma/K)^{(\times p)} = \langle Ka, Kb \rangle^{(\times p)} \cong C_p^{(\times 2p)}.$$

From (13), the image of x_i in $(\Gamma/K)^{(\times p)}$ is $v_i = (\mathbf{i} + \mathbf{1})(Ka)\mathbf{i}(Kb)^{a^{-1}}$. For all $1 \leq i \leq p-2$, v_i has non-zero Ka -component in its $(i+2)$ th entry, but zero Ka -component in its j th entry for all $j \geq i+3$, so that v_i is linearly independent of $\langle v_1, \dots, v_{i-1} \rangle$.

Similarly, for all $1 \leq i \leq p-2$, v_i has zero Kb -component in its $(p-1)$ th entry, whereas v_{p-1} has non-zero Kb -component there, so v_{p-1} is independent of $\langle v_1, \dots, v_{p-2} \rangle$. \square

Corollary 4.1. *For all $n \geq 1$,*

$$|\Gamma : \text{Stab}_\Gamma(n)| \geq p^{(p-2)(p^{n-1}-1)+1}.$$

Proof. It suffices to check $|\text{Stab}_\Gamma(m) : \text{Stab}_\Gamma(m+1)| \geq p^{(p-2)(p^{m-1}-1)}$ for all $m \geq 1$.

Embedding $\text{Stab}_\Gamma(m)$ into $\Gamma^{(\times p^m)}$ (via repeated application of ψ) we have an induced embedding:

$$\text{Stab}_\Gamma(m)/\text{Stab}_\Gamma(m+1) \hookrightarrow (\Gamma/\text{Stab}_\Gamma(1))^{(\times p^m)} \cong C_p^{(\times p^m)}$$

(with each $(\Gamma/\text{Stab}_\Gamma(1))$ -factor generated by $\text{Stab}_\Gamma(1)a$). From (13), the image of $x_i \in \text{Stab}_\Gamma(1)$ in $(\Gamma/\text{Stab}_\Gamma(1))^{(\times p)}$ is $(\mathbf{i} + \mathbf{1})(\text{Stab}_\Gamma(1)a)$.

Arguing as in the proof of Corollary 2.11 (ii), the elements:

$$\left(\prod_{j=1}^{p-2} x_j^{\lambda_{i,j}} \right)_{i=1}^{p^m}$$

are distinct modulo $\text{Stab}_\Gamma(m+1)$ as the $p^{m-1} \times (p-2)$ coefficients $\lambda_{i,j}$ vary over $\{0, 1, \dots, p-1\}$, and the required result follows. \square

We also introduce some further normal subgroups which will be useful “placeholders” for our induction in the proof of Theorem 1.4. For $1 \leq i \leq p$, let:

$$L_i = \langle x_i, \dots, x_{p-1}, K^{(\times p)} \rangle \leq K$$

(with $L_p = K^{(\times p)}$ by convention). For $0 \leq i \leq p$, let:

$$\begin{aligned} K_{\mathbf{i}}^{(\times p)} &= \langle \mathbf{i}(x_1), \dots, (\mathbf{p} - \mathbf{1})(x_1), L_2^{(\times p)} \rangle \leq K^{(\times p)}; \\ K_{\mathbf{i}}^{(\times p^{n+1})} &= (K_{\mathbf{i}}^{(\times p)})^{(\times p^n)} \end{aligned}$$

(with the convention $K_{\mathbf{p}}^{(\times p)} = L_2^{(\times p)}$). We thus have descending chains of subgroups:

$$L_2^{(\times p^n)} = K_{\mathbf{p}}^{(\times p^n)} \leq K_{\mathbf{p}-\mathbf{1}}^{(\times p^n)} \leq \dots \leq K_{\mathbf{1}}^{(\times p^n)} \leq K_{\mathbf{0}}^{(\times p^n)} = K^{(\times p^n)}$$

and:

$$K^{(\times p^{n+1})} = L_p^{(\times p^n)} \leq L_{p-1}^{(\times p^n)} \leq \dots \leq L_2^{(\times p^n)}.$$

Lemma 4.2. *The following are normal in $\Gamma_{(p)}$, for all $n \in \mathbb{N}$:*

- (i) $K^{(\times p^n)}$;
- (ii) $L_i^{(\times p^n)}$, for $1 \leq i \leq p-1$;
- (iii) $K_{\mathbf{i}}^{(\times p^{n+1})}$, for $0 \leq i \leq p-1$.

Proof. First, by induction we reduce to the case $n = 0$. For let $H = K^{(\times p^n)}$, $K_{\mathbf{i}}^{(\times p^{n+1})}$ or $L_i^{(\times p^n)}$. If $H \triangleleft \Gamma_{(p)}$, then $H^{(\times p)}$ is normalised by a (which permutes the factors) and by b (which acts on each factor as $a^{\pm 1}$ or b).

Now certainly $K \triangleleft \Gamma_{(p)}$, so we have (i). In the other cases, we check that conjugates of a generating set for the subgroup, by the generators of $\Gamma_{(p)}$, lie in the subgroup.

For (ii), consider the conjugation action of $\Gamma_{(p)}$ on $K/K^{(\times p)}$. We have $x_i^a = x_i x_{i+1}^{-1}$ for $1 \leq i \leq p-2$; $x_{p-1}^a \equiv x_{p-1} \pmod{K^{(\times p)}}$ and $x_i^b \equiv x_i \pmod{K^{(\times p)}}$ for $1 \leq i \leq p-1$. Thus $L_i \triangleleft \Gamma_{(p)}$.

For (iii), consider the conjugation action of $\Gamma_{(p)}$ on $K^{(\times p)}/L_2^{(\times p)}$. We have $\mathbf{i}(x_1)^a = \mathbf{i}(x_1)(\mathbf{i} + \mathbf{1})(x_1)^{-1}$ for $1 \leq i \leq p-2$; $(\mathbf{p} - \mathbf{1})(x_1)^a \equiv (\mathbf{p} - \mathbf{1})(x_1) \pmod{L_2^{(\times p)}}$ and $\mathbf{i}(x_1)^b \equiv \mathbf{i}(x_1)$ for $1 \leq i \leq p-1$. Thus $K_{\mathbf{i}}^{(\times p)} \triangleleft \Gamma_{(p)}$. \square

To apply Lemma 2.1 in our induction, we will need certain commutators of elements in Γ to lie in a sufficiently deep subgroup. As such, we note the following.

Lemma 4.3. (i) $[\text{Stab}_{\Gamma}(1), \text{Stab}_{\Gamma}(1)] \leq K^{(\times p)}$;

(ii) $[\text{Stab}_{\Gamma}(1), K^{(\times p)}] \leq L_2^{(\times p)}$.

Proof. For (i), note that any element of $\text{Stab}_{\Gamma}(1)$ is a p -tuple of elements of Γ , so any commutator of such is a p -tuple of elements of $[\Gamma, \Gamma] = K$.

Likewise in (ii), every element of $[\text{Stab}_{\Gamma}(1), K^{(\times p)}]$ is a p -tuple of elements of $[\Gamma, K]$. K/L_2 is generated by x_1 , and Γ is generated by a and b , so, since $L_2 \triangleleft K$, it suffices to check that $[a, x_1], [b, x_1] \in L_2$ (the former by definition; the latter by (i)). \square

We now describe our approximations to elements of Γ by commutators. These are split between the next four propositions, which are closely reminiscent of Proposition 3.6 in our proof for Grigorchuk's group. The first of these allows us to approximate elements of $K^{(\times p^2)}$ up to an error in $K_1^{(\times p^2)}$. We remark that this is the only point in the proof of Theorem 1.4 at which different arguments are required according to the value of p . To be precise, we exhibit one construction of an approximation by commutators which is valid for $p \geq 7$; one for $p = 5$ and one for $p = 3$.

Let $\mathfrak{c} : \Gamma \times \Gamma \rightarrow \Gamma$ be given by $\mathfrak{c}(g, h) = [g, h]$.

Proposition 4.4. *The restriction of \mathfrak{c} to $L_{p-1} \times K^{(\times p)}$ descends to a well-defined map:*

$$\bar{\mathfrak{c}}_0 : (L_{p-1}/K^{(\times p^2)}) \times (K^{(\times p)}/K^{(\times p^2)}) \rightarrow \Gamma/K_1^{(\times p^2)}$$

whose image contains $K^{(\times p^2)}/K_1^{(\times p^2)}$.

Proof. Observe that by Lemma 2.1, for $g_1 \in L_{p-1}$, $h_1 \in K^{(\times p)}$ and $g_2, h_2 \in K^{(\times p^2)}$ we have:

$$[g_1 g_2, h_1 h_2][g_1, h_1]^{-1} \in [L_{p-1}, K^{(\times p^2)}].$$

But L_{p-1} is generated by $K^{(\times p)}$ and x_{p-1} , both of which lie in $\text{Stab}_\Gamma(1)^{(\times p)}$ (the latter since x_{p-1} is congruent modulo $K^{(\times p)}$ to (b, \dots, b)). Hence:

$$[L_{p-1}, K^{(\times p^2)}] \subseteq [\text{Stab}_\Gamma(1), K^{(\times p)}]^{(\times p)} \subseteq L_2^{(p^2)} \subseteq K_1^{(\times p^2)} \text{ (by Lemma 4.3 (ii)).}$$

Thus $\bar{\mathfrak{c}}_0$ is indeed well-defined.

We now establish that the image of $\bar{\mathfrak{c}}_0$ contains $K^{(\times p^2)}/K_1^{(\times p^2)}$. First suppose $p \geq 7$. We have:

$$\begin{aligned} x_1^a &= (b, b^{-1}a, a^{-2}, a, 1, \dots, 1) \\ x_1^{a^{-2}} &= (a, 1, \dots, 1, b, b^{-1}a, a^{-2}) \end{aligned}$$

so that $\mathbf{0}(x_1) = [x_1^{a^{-2}}, x_1^a]$. Moreover for $\lambda \in \mathbb{N}$,

$$\mathbf{0}(x_1)^\lambda \equiv [x_1^{a^{-2}}, (x_1^a)^\lambda] \pmod{[[K, K]K]} \leq L_2^{(\times p)} \leq K_1^{(\times p)}$$

(by Lemma 4.3).

Now, any element of $K^{(\times p^2)}/K_1^{(\times p^2)}$ is represented by a vector $(\mathbf{0}(x_1)^{\lambda_j})_{j=1}^p$ for some $(\lambda_j)_{j=1}^p \in \mathbb{F}_p^p$. By the above,

$$(\mathbf{0}(x_1)^{\lambda_j})_{j=1}^p \equiv [(x_1^{a^{-2}})_{j=1}^p, ((x_1^a)^{\lambda_j})_{j=1}^p] \pmod{K_1^{(\times p^2)}}$$

and we are done.

Now suppose $p = 5$. We have $x_1^a = (b, b^{-1}a, a^{-2}, a, 1)$. Hence:

$$[x_1, x_1^a] = (x_1, aba^{-2}b^{-1}a, 1, 1, 1).$$

An easy direct computation yields $aba^{-2}b^{-1}a \equiv x_1 \pmod{L_2}$. So:

$$[x_1, x_1^a] \equiv \mathbf{0}(x_1)^2 \pmod{K_1^{(\times 5)}}.$$

As before, any element of $K^{(\times 5^2)}/K_1^{(\times 5^2)}$ is represented by:

$$(\mathbf{0}(x_1)^{\lambda_j})_{j=1}^5 \equiv [(x_1)_{j=1}^5, ((x_1^a)^{3\lambda_j})_{j=1}^5] \pmod{K_1^{(\times 5^2)}}$$

for some $(\lambda_j)_{j=1}^5 \in \mathbb{F}_5^5$, as required.

Finally suppose $p = 3$. Recall that:

$$b = (a, a^{-1}, b) \text{ and } x_1 = [a, b] = (b^{-1}a, a, ab).$$

Thus:

$$[b, x_1] = ([a, b^{-1}a], 1, [b, ab])$$

We compute:

$$\begin{aligned} [a, b^{-1}a] &= ([b, a]^{b^{-1}})^{a^{-1}} \\ &\equiv (x_1^{-1})^{a^{-1}} \pmod{K^{(\times 3)}} \text{ (by Lemma 4.3 (i))} \\ &\equiv x_1^{-1} \pmod{L_2} \end{aligned}$$

and:

$$\begin{aligned} [b, ab] &= [b, a][[b, a], b] \\ &\equiv x_1^{-1} \pmod{K^{(\times 3)}} \text{ (by Lemma 4.3 (i))} \end{aligned}$$

so that:

$$[b, x_1] \equiv (x_1^{-1}, 1, x_1^{-1}) \pmod{L_2^{(\times 3)}} \equiv \mathbf{0}(x_1) \pmod{K_1^{(\times 3)}}. \quad (16)$$

Now:

$$\begin{aligned} x_2 = [a, x_1] &= (b^{-1}a^{-1}b^{-1}a, a^{-1}ba, b) \\ &= (bx_1[x_1, b^{-1}], bx_1^{-1}, b) \\ &\equiv (bx_1, bx_1^{-1}, b) \pmod{K^{(\times 3^2)}} \text{ (by Lemma 4.3 (i))} \end{aligned}$$

so for $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{N}$,

$$[x_2, (x_1^{\lambda_1}, x_1^{\lambda_2}, x_1^{\lambda_3})] \equiv ([bx_1, x_1^{\lambda_1}], [bx_1^{-1}, x_1^{\lambda_2}], [b, x_1^{\lambda_3}]) \pmod{K_1^{(\times 3^2)}}$$

(by the well-definedness of $\bar{\tau}_0$) whereas for $\mu, \lambda \in \mathbb{N}$,

$$\begin{aligned} [bx_1^\mu, x_1^\lambda] &= [b, x_1^\lambda][[b, x_1^\lambda], x_1^\mu] \\ &\equiv [b, x_1^\lambda] \pmod{L_2^{(\times 3)}} \\ &\equiv [b, x_1]^\lambda \pmod{L_2^{(\times 3)}} \text{ (by Lemma 4.3 (ii))} \end{aligned}$$

so by (16) we have:

$$[x_2, (x_1^{\lambda_1}, x_1^{\lambda_2}, x_1^{\lambda_3})] \equiv (\mathbf{0}(x_1)^{\lambda_1}, \mathbf{0}(x_1)^{\lambda_2}, \mathbf{0}(x_1)^{\lambda_3}) \pmod{L_2^{(\times 3^2)}}$$

and every element of $K^{(\times p^2)}/K_{\mathbf{1}}^{(\times p^2)}$ is represented by $(x_1^{\lambda_1}, x_1^{\lambda_2}, x_1^{\lambda_3})$ for some $\lambda_i \in \mathbb{F}_3$, as required. \square

Second, we construct an approximation to elements of $K_{\mathbf{i}+1}^{(\times p^2)}$ up to an error lying in $K_{\mathbf{i}+2}^{(\times p^2)}$.

Proposition 4.5. *Let $0 \leq i \leq p-2$. The restriction of \mathfrak{c} to $L_{p-2} \times K_{\mathbf{i}}^{(\times p^2)}$ descends to a well-defined map:*

$$\bar{\mathfrak{c}}_i : (L_{p-2}/K^{(\times p^2)}) \times (K_{\mathbf{i}}^{(\times p^2)}/K_{\mathbf{i}+1}^{(\times p^2)}) \rightarrow \Gamma/K_{\mathbf{i}+2}^{(\times p^2)}$$

whose image contains $K_{\mathbf{i}+1}^{(\times p^2)}/K_{\mathbf{i}+2}^{(\times p^2)}$.

Proof. We check first that $\bar{\mathfrak{c}}_i$ is well-defined. Observe that by Lemma 2.1, for $g_1 \in L_{p-2}$, $h_1 \in K_{\mathbf{i}}^{(\times p^2)}$, $g_2 \in K^{(\times p^2)}$ and $h_2 \in K_{\mathbf{i}+1}^{(\times p^2)}$, we have:

$$[g_1 g_2, h_1 h_2][g_1, h_1]^{-1} \in [L_{p-2}, K_{\mathbf{i}+1}^{(\times p^2)}][K_{\mathbf{i}}^{(\times p^2)}, K^{(\times p^2)}].$$

It therefore suffices to check that:

$$[L_{p-2}, K_{\mathbf{i}+1}^{(\times p^2)}][K_{\mathbf{i}}^{(\times p^2)}, K^{(\times p^2)}] \subseteq K_{\mathbf{i}+2}^{(\times p^2)}. \quad (17)$$

But $L_{p-2} = \langle x_{p-2}, L_{p-1} \rangle$ and $L_{p-1} \subseteq \text{Stab}_{\Gamma}(2)$, so that by Lemma 4.3 (i),

$$[L_{p-1}, K_{\mathbf{i}+1}^{(\times p^2)}] \subseteq K^{(\times p^3)} \subseteq K_{\mathbf{i}+2}^{(\times p^2)}$$

while $x_{p-2}(a, \dots, a)^{-1} \in \text{Stab}_{\Gamma}(1)^{(\times p)}$, so:

$$[\langle x_{p-2} \rangle, K_{\mathbf{i}+1}^{(\times p^2)}] \subseteq [\langle a \rangle, K_{\mathbf{i}+1}^{(\times p)}]^{(\times p)} [\text{Stab}_{\Gamma}(1), K_{\mathbf{i}+1}^{(\times p)}]^{(\times p)} \subseteq K_{\mathbf{i}+2}^{(\times p^2)}$$

(using Lemma 4.3 (ii)). Thus $[L_{p-2}, K_{\mathbf{i}+1}^{(\times p^2)}] \subseteq K_{\mathbf{i}+2}^{(\times p^2)}$. Also, $K_{\mathbf{i}}^{(\times p^2)}$, $K^{(\times p^2)} \subseteq \text{Stab}_{\Gamma}(3)$, so by Lemma 4.3 (i),

$$[K_{\mathbf{i}}^{(\times p^2)}, K^{(\times p^2)}] \subseteq K^{(\times p^3)} \subseteq K_{\mathbf{i}+2}^{(\times p^2)}.$$

Thus (17) is indeed satisfied, and $\bar{\mathfrak{c}}_i$ is well-defined.

We now check that the image of $\bar{\mathfrak{c}}_i$ contains $K_{\mathbf{i}+1}^{(\times p^2)}/K_{\mathbf{i}+2}^{(\times p^2)}$. First note that for any $\lambda \in \mathbb{N}$,

$$\begin{aligned} (\mathbf{i} + \mathbf{1})(x_1)^{\lambda} &= (\mathbf{i} + \mathbf{1})(x_1^{\lambda}) \\ &= [a, \mathbf{i}(x_1^{\lambda})] \\ &= [a, \mathbf{i}(x_1)^{\lambda}]. \end{aligned}$$

Now, every element of $K_{\mathbf{i}+1}^{(\times p^2)}/K_{\mathbf{i}+2}^{(\times p^2)}$ is represented by an element:

$$((\mathbf{i} + \mathbf{1})(x_1)^{\lambda_j})_{j=1}^p = [(a)_{j=1}^p, (\mathbf{i}(x_1)^{\lambda_j})_{j=1}^p] \quad (18)$$

for some $(\lambda_j)_{j=1}^p \in \mathbb{N}^p$. From (14), there exist $y_1, \dots, y_p \in \text{Stab}_{\Gamma}(1)$ such that:

$$x_{p-2} = (ay_j)_{j=1}^p.$$

For $1 \leq j \leq p$,

$$[ay_j, \mathbf{i}(x_1)^{\lambda_j}][a, \mathbf{i}(x_1)^{\lambda_j}]^{-1} \in [\text{Stab}_\Gamma(1), K_{\mathbf{i}}^{(\times p)}] \subseteq K_{\mathbf{i}+\mathbf{2}}^{(\times p)}$$

by Lemma 2.1 and Lemma 4.3 (ii). Combining with (18) we have:

$$((\mathbf{i} + \mathbf{1})(x_1)^{\lambda_j})_{j=1}^p \equiv \mathbf{c}(x_{p-2}, (\mathbf{i}(x_1)^{\lambda_j})_{j=1}^p) \pmod{K_{\mathbf{i}+\mathbf{2}}^{(\times p^2)}}.$$

Since $x_{p-2} \in L_{p-2}$ and $(\mathbf{i}(x_1)^{\lambda_j})_{j=1}^p \in K_{\mathbf{i}}^{(\times p^2)}$ we are done. \square

Finally, we construct an approximation to elements of $L_{i+1}^{(\times p^2)}$ up to an error lying in $L_{i+2}^{(\times p^2)}$.

Proposition 4.6. *Let $1 \leq i \leq p-2$. The restriction of \mathbf{c} to $L_{p-2}^{(\times p)} \times L_i^{(\times p^2)}$ descends to a well-defined map:*

$$\bar{\mathbf{c}}_i : (L_{p-2}^{(\times p)} / K^{(\times p^2)}) \times (L_i^{(\times p^2)} / L_{i+1}^{(\times p^2)}) \rightarrow \Gamma / L_{i+2}^{(\times p^2)}$$

whose image contains $L_{i+1}^{(\times p^2)} / L_{i+2}^{(\times p^2)}$.

Proof. We check first that $\bar{\mathbf{c}}_i$ is well-defined. Observe that by Lemma 2.1, for $g_1 \in L_{p-2}^{(\times p)}$, $h_1 \in L_i^{(\times p^2)}$, $g_2 \in K^{(\times p^2)}$ and $h_2 \in L_{i+1}^{(\times p^2)}$, we have:

$$[g_1 g_2, h_1 h_2][g_1, h_1]^{-1} \in [L_{p-2}^{(\times p)}, L_{i+1}^{(\times p^2)}][K^{(\times p^2)}, L_i^{(\times p^2)}].$$

It therefore suffices to check that:

$$[L_{p-2}^{(\times p)}, L_{i+1}^{(\times p^2)}][K^{(\times p^2)}, L_i^{(\times p^2)}] \subseteq L_{i+2}^{(\times p^2)}. \quad (19)$$

Certainly, $[K^{(\times p^2)}, L_i^{(\times p^2)}] \leq K^{(\times p^3)} \leq L_{i+2}^{(\times p^2)}$ (by Lemma 4.3 (i)). Meanwhile,

$$[L_{p-2}^{(\times p)}, L_{i+1}^{(\times p^2)}] \subseteq [L_{p-2}, L_{i+1}^{(\times p)}]^{(\times p)}$$

so it suffices to check that $[L_{p-2}, L_{i+1}^{(\times p)}] \subseteq L_{i+2}^{(\times p)}$.

L_{p-2} is generated by x_{p-2} and $L_{p-1} \subseteq \text{Stab}_\Gamma(1)^{(\times p)}$ (the latter inclusion holds because L_{p-1} is generated by $K^{(\times p)}$ and $x_{p-1} \in (b, \dots, b)K^{(\times p)}$). Thus:

$$[L_{p-1}, L_{i+1}^{(\times p)}] \subseteq [\text{Stab}_\Gamma(1), L_{i+1}]^{(\times p)} \leq K^{(\times p^2)} \leq L_{i+1}^{(\times p)}.$$

Meanwhile, $L_{i+1}^{(\times p)} / L_{i+2}^{(\times p)}$ is generated by the images of $\mathbf{0}(x_{i+1}), \dots, (\mathbf{p}-\mathbf{1})(x_{i+1})$, so by (14),

$$[\langle x_{p-2} \rangle, L_{i+1}^{(\times p)}] \subseteq [\langle a \rangle, \langle x_{i+1} \rangle]^{(\times p)} L_{i+2}^{(\times p)} \subseteq L_{i+2}^{(\times p)}.$$

Thus (19) is indeed satisfied, and $\bar{\mathbf{c}}_i$ is well-defined.

We now check that the image of $\bar{\mathbf{c}}_i$ contains $L_{i+1}^{(\times p^2)} / L_{i+2}^{(\times p^2)}$.

First, for any $\lambda \in \mathbb{N}$,

$$x_{i+1}^\lambda = [a, x_i]^\lambda \equiv [a, x_i^\lambda] \pmod{K^{(\times p)}} \text{ (by Lemma 4.3 (i)).}$$

As in the proof of Proposition 4.5, there exist, by (14), $y_1, \dots, y_p \in \text{Stab}_\Gamma(1)$ such that:

$$x_{p-2} = (ay_j)_{j=1}^p.$$

For $1 \leq j \leq p$,

$$[ay_j, x_i^{\lambda_j}][a, x_i^{\lambda_j}]^{-1} \in [K, K] \leq K^{(\times p)}$$

(by Lemma 4.3 (i)). Thus for $\lambda_1, \dots, \lambda_p \in \mathbb{N}$,

$$[x_{p-2}, (x_i^{\lambda_1}, \dots, x_i^{\lambda_p})] \equiv (x_{i+1}^{\lambda_1}, \dots, x_{i+1}^{\lambda_p}) \pmod{K^{(\times p^2)}}.$$

Now every element of $L_{i+1}^{(\times p^2)} / L_{i+2}^{(\times p^2)}$ is represented by:

$$(x_{i+1}^{\lambda_1}, \dots, x_{i+1}^{\lambda_{p^2}})$$

for some $(\lambda_j)_{j=1}^{p^2} \in \mathbb{N}^{p^2}$. From the above,

$$(x_{i+1}^{\lambda_1}, \dots, x_{i+1}^{\lambda_{p^2}}) \equiv [(x_{p-2})_{j=1}^p, (x_i^{\lambda_j})_{j=1}^{p^2}] \pmod{K^{(\times p^3)}} \leq L_{i+2}^{(\times p^2)}$$

and we have $(x_{p-2})_{j=1}^p \in L_{p-2}^{(\times p)}$, $(x_i^{\lambda_j})_{j=1}^{p^2} \in L_i^{(\times p^2)}$ so the result follows. \square

Up to now, we have concentrated on approximating, by commutators, elements lying in $K^{(\times p^2)}$ but outside $K^{(\times p^3)}$. We can however quickly extend these approximations to elements lying between $K^{(\times p^m)}$ and $K^{(\times p^{m+1})}$ for arbitrary $m \geq 2$. Indeed, since the conclusions of Propositions 4.4-4.6 concern only computations within the group $K/K^{(\times p^3)}$ the generalisation from the case $m = 2$ is immediate from the identification:

$$K^{(\times p^{m-2})} / K^{(\times p^{m+1})} \cong (K / K^{(\times p^3)})^{(\times p^{m-2})}$$

and the observation that, in a direct product of groups, a tuple of commutators of elements of the factors is the commutator of the tuples of those same elements.

Proposition 4.7. *Let $m \geq 2$.*

- (i) *The restriction of \mathfrak{c} to $L_{p-1}^{(\times p^{m-2})} \times K^{(\times p^{m-1})}$ descends to a well-defined map:*

$$\bar{\mathfrak{c}}_{0,m} : (L_{p-1}^{(\times p^{m-2})} / K^{(\times p^m)}) \times (K^{(\times p^{m-1})} / K^{(\times p^m)}) \rightarrow \Gamma / K_1^{(\times p^m)}$$

whose image contains $K^{(\times p^m)} / K_1^{(\times p^m)}$.

- (ii) *Let $0 \leq i \leq p-2$. The restriction of \mathfrak{c} to $L_{p-2}^{(\times p^{m-2})} \times K_1^{(\times p^m)}$ descends to a well-defined map:*

$$\bar{\mathfrak{c}}_{i,m} : (L_{p-2}^{(\times p^{m-2})} / K^{(\times p^m)}) \times (K_1^{(\times p^m)} / K_{i+1}^{(\times p^m)}) \rightarrow \Gamma / K_{i+2}^{(\times p^m)}$$

whose image contains $K_{i+1}^{(\times p^m)} / K_{i+2}^{(\times p^m)}$.

(iii) Let $1 \leq i \leq p-2$. The restriction of \mathfrak{c} to $L_{p-2}^{(\times p^{m-1})} \times L_i^{(\times p^{m-2})}$ descends to a well-defined map:

$$\bar{\mathfrak{c}}_{i,m} : (L_{p-2}^{(\times p^{m-1})} / K^{(\times p^m)}) \times (L_i^{(\times p^m)} / L_{i+1}^{(\times p^m)}) \rightarrow \Gamma / L_{i+2}^{(\times p^m)}$$

whose image contains $L_{i+1}^{(\times p^m)} / L_{i+2}^{(\times p^m)}$.

Proof. By the preceding discussion, (i), (ii) and (iii) follow, respectively, from Propositions 4.4, 4.5 and 4.6. \square

We are now ready to put everything together, and use the approximations from Proposition 4.7 (i)-(iii) to prove a result closely analogous to Proposition 3.8: namely, if a symmetric subset $X \subseteq \Gamma$ contains an approximation to every element of Γ up to an error lying in $K^{(\times p^m)}$, then every element of Γ is approximated up to an error lying in $K^{(\times p^{m+1})}$ by a short word in X .

Proposition 4.8. *Let C_p be as in Theorem 1.3. Let $m \geq 2$ and let $X \subseteq \Gamma$ be a symmetric subset such that:*

$$XK^{(\times p^m)} = \Gamma. \quad (20)$$

Then:

$$X^{C_p} K^{(\times p^{m+1})} = \Gamma. \quad (21)$$

Proof. The first step shall be to show that:

$$K^{(\times p^m)} \subseteq X^4 K_1^{(\times p^m)}. \quad (22)$$

To this end let $k \in K^{(\times p^m)}$. By Proposition 4.7 (i), there exist $g \in L_{p-1}^{(\times p^{m-2})}$, $h \in K^{(\times p^{m-1})}$ such that:

$$[g, h] \equiv k \pmod{K_1^{(\times p^m)}}.$$

From (20), there exist $x_g, x_h \in X$ such that:

$$x_g \equiv g, x_h \equiv h \pmod{K^{(\times p^m)}}$$

so by the well-definedness of the map $\bar{\mathfrak{c}}_{0,m}$ from Proposition 4.7 (i),

$$X^4 \ni [x_g, x_h] \equiv k \pmod{K_1^{(\times p^m)}}$$

and we have (22).

Define the integer sequence $(a_n)_n$ recursively by $a_0 = 4$ and $a_n = 2a_{n-1} + 2$ for $n \geq 1$. The second step of the proof shall be to show that:

$$K_i^{(\times p^m)} \subseteq X^{a_i} K_{i+1}^{(\times p^m)} \quad (23)$$

for $0 \leq i \leq p-1$. This shall be achieved by induction on i , using Proposition 4.7 (ii) at each stage (and the base case $i = 0$ being provided by (22)).

For let $1 \leq i \leq p-1$ and let $k \in K_i^{(\times p^m)}$. By Proposition 4.7 (ii), there exist $g \in L_{p-2}^{(\times p^{m-2})}$, $h \in K_{i-1}^{(\times p^m)}$ such that:

$$[g, h] \equiv k \pmod{K_{i+1}^{(\times p^m)}}.$$

From (20) and the induction hypothesis, there exist $x_g \in X$, $x_h \in X^{a_{i-1}}$ such that:

$$x_g \equiv g \pmod{K^{(\times p^m)}}, x_h \equiv h \pmod{K_i^{(\times p^m)}}$$

so by the well-definedness of the map $\bar{c}_{i,m}$ from Proposition 4.7 (ii),

$$X^{a_i} = X^{2a_{i-1}+2} \ni [x_g, x_h] \equiv k \pmod{K_{i+1}^{(\times p^m)}}$$

and we have (23).

Define the integer sequence $(b_n)_n$ recursively by $b_1 = \sum_{n=0}^{p-1} a_n$ and $b_{n+1} = 2b_n + 2$ for $n \geq 2$. Combining the inclusions (23) for i from 0 to $p-1$, we have:

$$K^{(\times p^m)} = K_0^{(\times p^m)} \subseteq X^{b_1} K_p^{(\times p^m)} = X^{b_1} L_2^{(\times p^m)}. \quad (24)$$

Our third objective shall be to show that:

$$L_i^{(\times p^m)} \subseteq X^{b_i} L_{i+1}^{(\times p^m)} \quad (25)$$

for $1 \leq i \leq p-1$. This again shall be by induction on i , using Proposition 4.7 (iii), the base case $i = 1$ being provided by (24).

Thus let $2 \leq i \leq p-1$ and let $k \in L_i^{(\times p^m)}$. By Proposition 4.7 (ii), there exist $g \in L_{p-2}^{(\times p^{m-1})}$, $h \in L_{i-1}^{(\times p^m)}$ such that:

$$[g, h] \equiv k \pmod{L_{i+1}^{(\times p^m)}}.$$

By (20) and the induction hypothesis, there exist $x_g \in X$, $x_h \in X^{b_{i-1}}$ such that:

$$x_g \equiv g \pmod{K^{(\times p^m)}}, x_h \equiv h \pmod{L_i^{(\times p^m)}}$$

so by the well-definedness of the map $\bar{c}_{i,m}$ from Proposition 4.7 (iii),

$$X^{b_i} = X^{2b_{i-1}+2} \ni [x_g, x_h] \equiv k \pmod{L_{i+1}^{(\times p^m)}}$$

as desired.

Finally, set $C_p = 1 + \sum_{i=1}^{p-1} b_i$. Expressing a_n and b_n in closed form, C_p is as in the statement of Theorems 1.3 and 1.4. We combine the inclusions (25) for i from 1 to $p-1$ to obtain:

$$K^{(\times p^m)} = L_1^{(\times p^m)} \subseteq X^{C_p-1} L_p^{(\times p^m)} = X^{C_p-1} K^{(\times p^{m+1})}.$$

Combining this last inclusion with (20), we have $\Gamma = X^{C_p} K^{(\times p^{m+1})}$, as required. \square

Proof of Theorem 1.4. Let $S \subseteq \Gamma/K^{(\times p^n)}$. If $n \leq 2$ then:

$$\text{diam}(\Gamma/K^{(\times p^n)}) \leq |G : K^{(\times p^2)}| = p^{p+1}.$$

If $n \geq 3$ then let $\tilde{S} \subseteq \Gamma$ be any subset whose image in $\Gamma/K^{(\times p^n)}$ is S . Then $B_{\tilde{S}}(p^{p+1})K^{(\times p^2)} = \Gamma$, and by repeated application of Proposition 4.8,

$$B_{\tilde{S}}(p^{p+1}C_p^{n-2})K^{(\times p^n)} = \Gamma.$$

Thus:

$$\text{diam}(\Gamma/K^{(\times p^n)}, S) \leq p^{p+1}C_p^{n-2} \ll_p p^{\log(C_p)n/\log(p)}$$

The result follows, since $|\Gamma : K^{(\times p^n)}| = p^{p^n+1}$. \square

Proof of Theorem 1.3. By Lemma 2.12, we have:

$$\text{diam}(\Gamma/\text{Stab}_\Gamma(n)) \leq \text{diam}(\Gamma/K^{(\times p^{n-1})}).$$

The result now follows from $|\Gamma : K^{(\times p^n)}| = p^{p^n+1}$; Theorem 1.4 and Corollary 4.1. \square

5 The Gupta-Sidki 3-Group

In this Section we deduce Theorem 1.5 from Theorem 1.4. We shall require some facts about the lower central series of $\Gamma = \Gamma_{(3)}$, which were established in [3]. Define the integer sequence (α_n) by $\alpha_1 = 1$, $\alpha_2 = 2$, $\alpha_n = 2\alpha_{n-1} + \alpha_{n-2}$ for $n \geq 3$, and set $\beta_n = \sum_{i=1}^n \alpha_i$. We have:

$$\begin{aligned} \alpha_n &= \frac{1}{2\sqrt{2}}((1 + \sqrt{2})^n - (1 - \sqrt{2})^n), \\ \beta_n &= \frac{1}{4}((1 + \sqrt{2})^{n+1} + (1 - \sqrt{2})^{n+1} - 2). \end{aligned}$$

Theorem 5.1 ([3]). *Let $X_1, \dots, X_n \in \{\mathbf{0}, \mathbf{1}, \mathbf{2}\}$. Then:*

$$\begin{aligned} \deg X_1 \cdots X_n(x_1) &= 1 + \sum_{i=1}^n X_i \alpha_i + \alpha_{n+1} \\ \deg X_1 \cdots X_n(x_2) &= 1 + \sum_{i=1}^n X_i \alpha_i + 2\alpha_{n+1}. \end{aligned}$$

Corollary 5.2. *For all $m \in \mathbb{N}$,*

$$K^{(\times 3^m)} \leq \gamma_{\alpha_{m+1}+1}(\Gamma) \leq \gamma_{\beta_{m+1}}(\Gamma).$$

Corollary 5.3. *For all $m \in \mathbb{N}$,*

$$|\Gamma : \gamma_{\beta_{m+1}}(\Gamma)| = 3^{(3^m+1)/2}.$$

Proof of Theorem 1.5. For $n = 1$ there is nothing to prove. Otherwise, let $m \in \mathbb{N}$ be such that $\beta_m + 1 \leq n \leq \beta_{m+1} + 1$. Then by Corollary 5.2,

$$\begin{aligned}
\text{diam}(\Gamma/\gamma_n(\Gamma)) &\leq \text{diam}(\Gamma/K^{(\times 3^{m+1})}) \\
&\ll 3^{\log(111)m/\log(3)} \quad (\text{by Theorem 1.4}) \\
&\ll n^{\log(111)/\log(1+\sqrt{2})} \\
&\ll (\log|\Gamma : \gamma_{\beta_m+1}(\Gamma)|)^{\log(111)/\log(3)} \quad (\text{by Corollary 5.3}) \\
&\leq (\log|\Gamma : \gamma_n(\Gamma)|)^{\log(111)/\log(3)}.
\end{aligned}$$

□

6 Spectral Gap and Mixing Time

For G a finite group and $S \subseteq G$ a symmetric subset, let A_S be the (normalized) adjacency operator on the Cayley graph $\text{Cay}(G, S)$. A_S is a self-adjoint operator of norm 1; let its spectrum be:

$$1 = \lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{|G|} \geq -1$$

with the eigenvalue $\lambda_1 = 1$ corresponding to the constant functionals on G . More generally, the 1-eigenspace of A_S is spanned by the indicator functions of the connected components of $\text{Cay}(G, S)$; in particular $1 > \lambda_2$ if and only if S generates G , and in this case the quantity $1 - \lambda_2$ is known as the *spectral gap* of the pair (G, S) .

The existence of a large spectral gap for a family of Cayley graphs is a matter of great interest. If a family of finite graphs of bounded valence with vertex sets of unbounded size possess a spectral gap bounded below by an absolute positive constant, then the graphs form an *expander family*. Expander graphs (and especially expander *Cayley* graphs) have multifarious applications across pure mathematics and theoretical computer science [25].

Now let Γ be \mathfrak{S} or $\Gamma_{(p)}$ and let $(\Gamma_i)_i$ be one of the descending sequences of finite-index normal subgroups from Theorems 1.1-1.5. Cayley graphs of the quotient groups Γ/Γ_i do not in general form expander families: for instance if S is a finite symmetric generating set for Γ , and S_i is the image of S in Γ/Γ_i , then the spectral gap of $\text{Cay}(\Gamma/\Gamma_i, S_i)$ tends to 0 as $i \rightarrow \infty$ (this follows from the fact that Γ is amenable [7] and Γ_i exhausts Γ). We do however have a weaker lower bound on the spectral gap of any connected Cayley graph of Γ/Γ_i , coming from our upper bounds on diameter and the following general inequality.

Proposition 6.1 ([11] Corollary 3.1). *Let G be a finite group and let S be a symmetric generating set. Then the spectral gap of (G, S) is $\geq (|S| \text{diam}(G, S)^2)^{-1}$.*

Theorems 1.1-1.5 combine with Proposition 6.1 to yield the following bounds on spectral gaps.

Corollary 6.2. *Let S be an arbitrary generating set for the finite group G . Denote by $\epsilon(G, S)$ the spectral gap of the pair (G, S) . Let C_p be as in Theorem 1.3.*

- (i) *If $G = \mathfrak{S}/\text{Stab}_{\mathfrak{S}}(n)$ then $\epsilon(G, S) = \Omega(|S|^{-1} \exp(-2 \log(35)n))$;*
- (ii) *If $G = \mathfrak{S}/\gamma_n(\mathfrak{S})$ then $\epsilon(G, S) = \Omega(|S|^{-1} n^{-2 \log(35)/\log(2)})$;*
- (iii) *If $G = \Gamma_{(p)}/\text{Stab}_{\Gamma_{(p)}}(n)$ then $\epsilon(G, S) = \Omega_p(|S|^{-1} \exp(-2 \log(C_p)n))$;*
- (iv) *If $G = \Gamma_{(p)}/K^{(\times p^n)}$ then $\epsilon(G, S) = \Omega_p(|S|^{-1} \exp(-2 \log(C_p)n))$;*
- (v) *If $G = \Gamma_{(3)}/\gamma_n(\Gamma_{(3)})$ then $\epsilon(G, S) = \Omega(|S|^{-1} n^{-2 \log(111)/\log(1+\sqrt{2})})$.*

A second closely related numerical invariant of finite Cayley graphs is the *mixing time*. This is a measure of the time taken for a lazy random walk on the Cayley graph to approach the uniform distribution. It may be defined as follows. Let $f_0 = \delta_e$ be the Dirac mass at the identity of G . The lazy random walk on $\text{Cay}(G, S)$ is defined by the operator $T_S = (A_S + I)/2$, where I is the identity operator on $\text{Cay}(G, S)$, and describes the progress on $\text{Cay}(G, S)$ of a particle which starts at the identity, and which at each step with equal probability either traverses an edge (chosen uniformly at random) or remains stationary. Recursively define $f_{l+1} = T_S(f_l)$, the distribution of the walk at time l . We may consider the walk to be well-mixed when f_l is close to the uniform distribution, in some appropriate norm on the complex functionals on G . Here we focus on mixing with respect to the ℓ^∞ -norm.

Definition 6.3. *Let G be a finite group and S be a symmetric generating set. The ℓ^∞ -mixing time of the pair (G, S) is the smallest positive integer l such that:*

$$\|f_l - \frac{1}{|G|} \chi_G\|_\infty \leq \frac{1}{2|G|}.$$

It may be easily seen that the LHS of the above inequality is a non-increasing function of l , so that once the random walk reaches its mixing time, it remains well-mixed thereafter. There is a close relationship between mixing time and spectral gap.

Proposition 6.4 ([24] Theorem 5.1). *Suppose the pair (G, S) has spectral gap $\epsilon > 0$. Then there exists an absolute constant $C > 0$ such that the ℓ^∞ -mixing time of (G, S) is at most $(C/\epsilon) \log|G|$.*

Applying Proposition to the conclusions of Corollary 6.2, we have corresponding bounds on mixing times, as follows.

Corollary 6.5. *Denote by $\mu(G, S)$ the ℓ^∞ -mixing time of the pair (G, S) .*

- (i) *If $G = \mathfrak{S}/\text{Stab}_{\mathfrak{S}}(n)$ then $\mu(G, S) = O(|S| \exp(\log(2450)n))$;*
- (ii) *If $G = \mathfrak{S}/\gamma_n(\mathfrak{S})$ then $\mu(G, S) = O(|S| n^{2 \log(35)/\log(2)+1})$;*

- (iii) If $G = \Gamma_{(p)} / \text{Stab}_{\Gamma_{(p)}}(n)$ then $\mu(G, S) = O_p(|S| \exp(\log(pC_p^2)n))$;
- (iv) If $G = \Gamma_{(p)} / K^{(\times p^n)}$ then $\mu(G, S) = O_p(|S| \exp(\log(pC_p^2)n))$;
- (v) If $G = \Gamma_{(3)} / \gamma_n(\Gamma_{(3)})$ then $\mu(G, S) = O(|S| n^{\log(36963)/\log(1+\sqrt{2})})$.

7 Growth in Branch Groups

Given a finitely generated group G and a finite generating set $S \subseteq G$, let $f_{(G,S)}(n) = |B_S(n)|$ be the *growth function*. Although for a given group G , the function $f_{(G,S)}$ may vary according to the generating set S , it only does so up to an appropriate notion of equivalence of functions. As such, we may speak without ambiguity about groups of *polynomial growth*, *exponential growth* and so on (see [10] Chapters VI-VII).

One of the key sources of interest in branch groups is the fact that they include many examples of groups with exotic growth behaviour. In particular, \mathfrak{G} has *intermediate growth*, that is: growth faster than any polynomial function but slower than any exponential function.

The following elementary fact exhibits a relationship between growth and diameter.

Lemma 7.1. *Let F be a finite group, and let $\phi : G \rightarrow F$ be an epimorphism. Then:*

$$f_{(G,S)}(\text{diam}(F, \phi(S))) \geq |F|. \quad (26)$$

This inequality suggests the following definition, which is made by analogy with that of the diameter of a finite group. Let $f_G(n)$ be the minimal value of $f_{(G,S)}(n)$, as S ranges over all finite generating subsets of G . From (26) we immediately obtain:

$$f_G(\text{diam}(F)) \geq |F|. \quad (27)$$

The relationship between growth and diameter can be exploited to yield information about both. For instance, from Theorems 1.2 and 1.4 we have the following.

Corollary 7.2. *There exist constants $\alpha(p) > 0$ such that:*

$$f_{\mathfrak{G}}(n) \gg \exp(\alpha(2)n^{\beta(\mathfrak{G})}) \text{ and } f_{\Gamma_{(p)}}(n) \gg_p \exp(\alpha(p)n^{\beta(\Gamma_{(p)})}),$$

where $\beta(\mathfrak{G}) = \log(2)/\log(35) \approx 0.195$ and $\beta(\Gamma_{(p)}) = \log(p)/\log(C_2(p))$ (here $C_2(p)$ is as in Theorem 1.3).

The bounds in Corollary 7.2 are not the best known: a slight modification of an argument of Grigorchuk [17] shows that if G is any finitely generated residually virtually nilpotent group, then either G is virtually nilpotent or $f_G(n) \geq \exp(\sqrt{n})$ (see [4]). In particular the latter conclusion applies to \mathfrak{G} and $\Gamma_{(p)}$. It is however possible that improvements upon the diameter bounds in Theorems 1.2 and 1.4 and their corollaries could yield new lower bounds on $f_{\mathfrak{G}}$ and $f_{\Gamma_{(p)}}$.

Conversely, known upper bounds on the growth translate into lower bounds on the diameters of finite quotients. In the case of \mathfrak{G} , the best upper bound on the growth is the following result of Bartholdi.

Theorem 7.3 ([16]). *Let $a, b, c, d \in \mathfrak{G}$ be as in subsection 2.2.1 and let $S = \{a, b, c, d\}$. Then:*

$$f_{(\mathfrak{G}, S)}(n) \ll \exp(n^\beta),$$

where $\beta = \frac{\log(2)}{\log(2) - \log(\eta)} \approx 0.768$, for η the real root of $X^3 + X^2 + X = 2$.

Corollary 7.4. *There exist absolute constants $C, C' > 0$ such that:*

$$\begin{aligned} \text{diam}(\mathfrak{G}/\text{Stab}_{\mathfrak{G}}(n)) &\geq C(\log|\mathfrak{G} : \text{Stab}_{\mathfrak{G}}(n)|)^{1/\beta}; \\ \text{diam}(\mathfrak{G}/\gamma_n(\mathfrak{G})) &\geq C'(\log|\mathfrak{G} : \gamma_n(\mathfrak{G})|)^{1/\beta} \end{aligned}$$

where β is as in Theorem 7.3.

Corollary 7.4 places a limit on the extent to which the constant $\log(35)/\log(2)$ appearing in Theorems 1.1 and 1.2 might be reduced (though it is almost certainly not sharp). It is unclear at this time whether the constant $1/\beta \approx 1.303$ in Corollary 7.4 could be close to sharp.

Acknowledgements

I would like to thank Laurent Bartholdi and Alejandra Garrido for illuminating discussions. Parts of this work were completed while the author was a fellow-commoner at Trinity Hall, Cambridge. I would like to thank the fellows of the College for providing me with a warm welcome and pleasant working conditions.

References

- [1] L. Babai, Á. Seress. *On the Diameter of Permutation Groups*. Europe J. Combinatorics, Vol. 13 (1992) 231–243
- [2] L. Bartholdi. *The Growth of Grigorchuk’s Group*. Internat. Math. Res. Notices 20 (1998) 1049–1054
- [3] L. Bartholdi. *Lie Algebras and Growth in Branch Groups*. Pacific Journal of Mathematics, Vol. 218, No. 1 (2005) 241–282
- [4] L. Bartholdi. *Growth of Groups and Wreath Products*. arXiv:1512.07044
- [5] L. Bartholdi, R. Grigorchuk. *Lie methods in growth of groups and groups of finite width*. In Computational and geometric aspects of modern algebra, London Math. Soc. Lecture Note Series 275, Cambridge University Press (2000), 1–27
- [6] L. Bartholdi, R. Grigorchuk, Z. Šuník. *Branch Groups*. Handbook of algebra, Vol. 3, North-Holland, Amsterdam (2003), 989–1112

- [7] L. Bartholdi, V.A. Kaimanovich, V.V. Nekrashevych. *On Amenability of Automata Groups*. Duke Math. J., 154 Issue 3 (2010) 575–598.
- [8] H. Bradford. *New Uniform Diameter Bounds in Pro- p Groups*. arXiv:1410.3007 [math.GR]
- [9] H. Bradford. *A Potent Solovay-Kitaev Procedure*. In preparation.
- [10] P. de la Harpe. *Topics in Geometric Group Theory*. University of Chicago Press (2000)
- [11] P. Diaconis, L. Saloff-Coste. *Comparison techniques for random walk on finite groups*. Ann. Probab. 21, Issue 4 (1993), 2131–2156.
- [12] O. Dinai. *Diameters of Chevalley groups over local rings*. Archiv der Mathematik 99, Issue 5 (2012) 417–424.
- [13] J. Driscoll, M. Furst. *Computing short generator sequences*. Information and Computation 72 (1987) 117–132
- [14] A. Gamburd, M. Shahshahani. *Uniform diameter bounds for some families of Cayley graphs*. Internat. Math. Res. Notices 71 (2004) 3813–3824
- [15] A. Garrido. *Abstract commensurability and the Gupta-Sidki group*. arXiv:1310.0493 [math.GR]
- [16] R. Grigorchuk. *On Burnside's problem on periodic groups*. Funktsional. Anal. i Prilozhen. 14 (1980), no. 1, 53–54, English translation: Functional Anal. Appl. 14 (1980), 41–43
- [17] R. Grigorchuk. *On the Hilbert-Poincaré series of graded algebras that are associated with groups*. Math. USSR-Sb. 66 (1990) no. 1, 211–229
- [18] R. Grigorchuk. *Just infinite branch groups*. In: New horizons in pro- p groups (Markus P. F. du Sautoy Dan Segal and Aner Shalev, eds.), Birkhäuser Boston, Boston, MA, 2000, pp. 121–179
- [19] N. Gupta, S. Sidki. *On the Burnside problem for periodic groups*. Math. Z., 182(3) (1983) 385–388
- [20] H.A. Helfgott. *Growth in groups: ideas and perspectives*. Bull. Amer. Math. Soc. 52 (2015), 357–413
- [21] H.A. Helfgott, Á. Seress. *On the diameter of permutation groups*. arXiv:1109.3550 (2011).
- [22] L. Kaloujnine. *La structure des p -groupes de Sylow des groupes symétriques finis*. Ann. École Norm. Sup. (3) 65 (1948), 239–276.
- [23] D. Kornhauser, G. Miller, P. Spirakis. *Coordinating pebble motion on graphs, the diameter of permutation groups, and applications*. MS thesis. M.I.T., Dept. of Electrical Engineering and Computer Science, 1984.

- [24] L. Lovász. *Random walks on graphs: a survey*. Combinatorics, Paul Erdős is Eighty 2 (1993), 1–46.
- [25] A. Lubotzky. Expander Graphs in Pure and Applied Mathematics. Bull. Amer. Math. Soc. 49, Issue 1 (2012), 113–162.
- [26] A.V. Rozhkov. *Lower central series of a group of tree automorphisms*. Mat. Zametki Vol. 60, No. 2 (1996) 225–237
- [27] S. Sidki. *On a 2-generated infinite 3-group: subgroups and automorphisms*. J. Algebra, 110 No. 1 (1987) 24–55
- [28] T. Tao. *Expansion in finite simple groups of Lie type*. AMS Graduate Studies in Mathematics 164, Providence, Rhode Island, 2015
- [29] J.S. Wilson. *Groups with every proper quotient finite*. Math. Proc. Cambridge Philos. Soc. 69 (1971), 373–391

H. Bradford, UNIVERSITÄT GÖTTINGEN, GERMANY
E-mail address: `henry.bradford@mathematik.uni-goettingen.de`